

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 28, 2004, 01:16:04 ; Search time 511 Seconds  
(without alignments)  
3649.824 Million cell updates/sec

Title: US-09-964-277-21

Perfect score: 517

Sequence: 1 MLPPLSLQTVFSLYFWNWR.....LKGVSQSSFGSMEIIIVS 517

Scoring table: OLIGO  
Xgapop 60.0, Xgapext 60.0  
Ygapop 60.0, Ygapext 60.0  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 2353733 seqs, 1803733377 residues

Word size: 1

Total number of hits satisfying chosen parameters: 4700994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-O=/cgn2\_1/USPTO.spool/US09964277/runat\_25022004.085402\_2242/app\_query.fasta\_1.711  
-DB=Published Applications NA -QWTF=fastap -SUFFIX=olip2n.rnnpb -MINMATCH=0.1  
-LOOPCL=0 -LOPEXT=0 -UNITS=bits -START=1 -END=1 -MAIRIX=oligo  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=quality -THR MIN=1  
-ALIGN=15 -MODE=LOCAL -OUTPMT=ptc -NORM=ext -HEAFSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US09964277@cgn 1.1.221 @runat\_25022004.085402\_2242  
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09D\_PUBCOMB.seq\*
- 13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq\*
- 16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq\*
- 17: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq\*
- 18: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
------------	-------	-------------	-----------	----	-------------

Alignment Scores:					
Pred. No.:	0				Length: 3332
Score:	517.00				Matches: 517
Percent Similarity:	100.00%				Conservative: 0
Best Local Similarity:	100.00%				Mismatches: 0

ALIGNMENTS

RESULT 1  
US-09-964-277-20  
; Sequence 20, Application US/09964277  
; Patent No. US20020137170A1  
; GENERAL INFORMATION:  
; APPLICANT: Luche, Ralf M.  
; APPLICANT: Wei, Bo  
; TITLE OF INVENTION: DSP-16 DUAL-SPECIFICITY PHOSPHATASE  
; FILE REFERENCE: 200125.434  
; CURRENT APPLICATION NUMBER: US/09/964,277  
; CURRENT FILING DATE: 2001-09-25  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 20  
; LENGTH: 3332  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-964-277-20

Sequence 20, Appl  
Sequence 3, Appl  
Sequence 27, Appl  
Sequence 673, App  
Sequence 1, Appl  
Sequence 1, Appl  
Sequence 25, Appl  
Sequence 26234, A  
Sequence 235, App  
Sequence 257, App  
Sequence 2429, App  
Sequence 8120, Ap  
Sequence 1, Appl  
Sequence 463, App  
Sequence 1814, Ap  
Sequence 1817, A  
Sequence 8068, Ap  
Sequence 31436, A  
Sequence 3118, Ap  
Sequence 3, Appl  
Sequence 69, Appl  
Sequence 69, Appl  
Sequence 49, Appl  
Sequence 11, Appl  
Sequence 265, App  
Sequence 654, App  
Sequence 270, App  
Sequence 1135, Ap  
Sequence 183, App  
Sequence 4, Appl  
Sequence 48, Appl  
Sequence 1, Appl  
Sequence 69, Appl  
Sequence 330, App  
Sequence 12, Appl  
Sequence 18, Appl  
Sequence 241, App  
Sequence 15, Appl  
Sequence 16818, A  
Sequence 310, App  
Sequence 310, App  
Sequence 310, App  
Sequence 310, App  
Sequence 310, App  
Sequence 310, App

Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0  
US-09-964-277-21 (1-517) x US-09-964-277-20 (1-3332)  
QY 1 MetLeuProLeuSerLeuGlnThrValPheSerLeuThrPheProValAsnThrArgArg 20  
DB 842 ATGTTGGCTCTCTCTCTTCCAGACTGTTTCTCAGTCTACTTCTGGGTAACTGGAGAGA 901  
QY 21 AlaserThrLeuPheThrCysLeuGlnGluMetGlnGlnAnGlyLeuGlyThrVal 40  
DB 902 GCITCAACTCTGTTCACTCTGCTGAGAGAGTGTGACAGAGATGGGATGGATATGTTG 961  
QY 41 LeuAsnAlaserAsnThrCysProLysProAspPheLeuProGluSerHisPheLeuArg 60  
DB 962 TTAATGCCAGCAATACCTGTCCAAAGCTGACATTATCCCGAGTCTCATTTCTGTGCT 1021  
QY 61 ValProValAsnAspSerPheCysGluLysIleLeuProThrLeuAspLysSerValAsp 80  
DB 1022 GTGCTGTGAATGACAGCTTTTGTGAGAAATTTTGGCGTGGTGGACAAATCAGTAGAT 1081  
QY 81 PheLeuGluLysAlaLysAlaserAsnGlyCysValLeuValHisCysLeuAlaGlyIle 100  
DB 1082 TTCAATGAGAAAGCAAAAGCCCTCAATGATGTGTTCTAGTGACATGTTTAGCTGGATC 1141  
QY 101 SerArgSerAlaThrIleAlaIleAlaThrIleMetLysArgMetAspMetSerLeuAsp 120  
DB 1142 TCCGCTCCGCCACCATCGTATCGCTACATCATGAAGAGATGGACATGCTTTAGAT 1201  
QY 121 GluAlaThrArgPheValLysGluLysArgProThrIleSerProAsnPheAsnPheLeu 140  
DB 1202 GAAGCTTACAGATTTGTGAAAGAAAGAAAGACCTATATCTCAAACTTCAATTTTCTG 1261  
QY 141 GlyLysLeuLeuAspThrGluLysLysIleLysAsnGlnThrGlyAlaSerGlyProLys 160  
DB 1262 GGCCAACTCTCGACTATCAGAGAGAGATTAAGAACACAGATGGAGCATCAGGGCCAAAG 1321  
QY 161 SerLysLeuLysLeuLysIleGluLysProAsnGluProValProAlaValSerGlu 180  
DB 1322 AGCAACTCAAGCTGTGCTGACCTTGAGAGAGCAAAATGAACCTGCTGCTGCTCAGAG 1381  
QY 181 GlyGlyLysSerGluThrProLeuSerProCysAlaAspSerAlaThrSerGlu 200  
DB 1382 GGTGGACAGAAAGCAGAGAGCGCCCTCACTGTCAGCTGCGCACTCTGTACCTCAGAG 1441  
QY 201 AlaAlaGlyGlnArgProValHisProAlaSerValProSerValProSerValGlnPro 220  
DB 1442 GCAGCAGCAAAAGGCCCTGTGATCCGCCAGCTGTCAGCTGCCAGCTGCCAGCG 1501  
QY 221 SerLeuLeuGluAspSerProLeuValGlnAlaLeuSerGlyLeuHisLeuSerAlaAsp 240  
DB 1502 TCGTGTGTAGAGACAGCCCGCTGTGTACAGCGCTCACTGGGCTGCACTGTCCGAGAC 1561  
QY 241 ArgLeuGluAspSerAsnLysLysLysArgSerPheSerLeuAspLysSerValSer 260  
DB 1562 AGGTGGAACAGACGAATTAAGCTCAAGCTTCTCTCTGGATATCAAACTCAGTTTCA 1621  
QY 261 TyrSerAlaSerMetAlaAlaSerLeuHisGlyPheSerSerSerGluAspAlaLeuGlu 280  
DB 1622 TATTCAGCCAGATGGCAGCATCTTACATGGCTTCTCCATCAGAGATGCTTTGGAA 1681  
QY 281 TyrThrLysProSerThrThrLeuAspGlyThrAsnLysLeuCysGlnPheSerProVal 300  
DB 1682 TACTACAACTTCCACTACTCTGGATGGAGCAACCAAGCTATGCCAGTTCTCCCTGTT 1741  
QY 301 GlnGluLeuSerGluGlnThrProGluThrSerProAspLysGluGluAlaSerIlePro 320  
DB 1742 CAGAACTATCGAGGAGAGACTCCGAAACCACTCTGTATAGAGAGAGCAGCATCCCT 1801  
QY 321 LysLysLeuGlnThrAlaArgProSerAspSerGlnSerLysArgLeuHisSerValArg 340  
DB 1802 AAGAAGTGACAGCCCGCCAGGCTTTCAGACAGCCAGAGCAAGGATGTCATTCGTCAGA 1861

QY 341 ThrSerSerSerGlyThrAlaGlnArgSerLeuLeuSerProLeuHisArgSerGlySer 360  
DB 1862 ACCAGCAGCAGTGGCACCGCCAGAGTCCCTTTATCTCCACTGCATCGAAGTGGAGC 1921  
QY 361 ValGluLeuAspLeuThrSerPheLeuPheGlyLeuSerThrSerGlnGlnHisLeu 380  
DB 1922 GTGGAGGAGCAATTAACACACAGCTTCCCTTTTGGGCTTTTCCACAGCAGCAGCCTC 1981  
QY 381 ThrLysSerAlaGlyLeuGlyLeuLysGlyTTPHisSerAspIleLeuAlaProGlnThr 400  
DB 1982 ACAGAGTCTGTGGCTGGCCCTTAAGGGCTGGCACTCGGATATCTTTGGCCCCCAGACC 2041  
QY 401 SerThrProSerLeuThrSerSerTrpPheAlaThrGluSerSerHisPheThrSer 420  
DB 2042 TCTACCCCTTCCCTGACACAGCTGTGTTATTTGCCACAGAGTCTCCACACTTCTACTCT 2101  
QY 421 AlaserAlaIleThrGlyGlySerAlaSerThrSerAlaThrSerCysSerGlnLeuPro 440  
DB 2102 GCCTCAGCATCTACGGAGCAGTGCAGTACTTCTGCTTACAGTGCAGCAGCTGCC 2161  
QY 441 ThrCysGlyAspGlnValThrSerValArgArgGlnLysProSerAspArgAlaAsp 460  
DB 2162 ACTTGGCAGAGCAACAGTCTATTCTGTGCGCAGCGCAGAGCAAGTGCAGAGCTGAC 2221  
QY 461 SerArgArgSerTrpHisGluSerProPheGluLysGlnPheLysArgArgSerCys 480  
DB 2222 TCGCGCGAGCTGGCATGAGAGAGCCCTTTTGAAGAGCTTTAAACGAGAGCTGC 2281  
QY 481 GlnMetGluPheGlyGluSerIleMetSerGluAsnArgSerArgGluLeuGlyLys 500  
DB 2282 CAATGTGAATTTGGAGAGAGCATCATGTGAGAGACAGGTCAACGAGAGAGCTGGGAAA 2341  
QY 501 ValGlySerGlnSerPheSerGlySerMetGluIleIleGluValSer 517  
DB 2342 GTGGSCAGTCACTAGCTTTTTCGGSCACATGGAATCATTGAGTCTCC 2392  
RESULT 2  
US-09-816-494-3  
; Sequence 3, Application US/09816494  
; Patent No. US20020034807A1  
; GENERAL INFORMATION:  
; APPLICANT: Meyers, Rachel A.  
; TITLE OF INVENTION: 38692 AND 21117, NOVEL DUAL SPECIFICITY  
; TITLE OF INVENTION: PHOSPHATASE MOLECULES AND USES THEREFOR  
; FILE REFERENCE: 10448-030002  
; CURRENT APPLICATION NUMBER: US/09/816,494  
; CURRENT FILING DATE: 2001-03-23  
; PRIOR APPLICATION NUMBER: US 60/191,858  
; PRIOR FILING DATE: 2000-03-24  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 1998  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-816-494-3  
Alignment Scores:  
Pred. No.: 0 Length: 1998  
Score: 488.00 Matches: 488  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 94.39% Indels: 0  
DB: 9 Gaps: 0  
US-09-964-277-21 (1-517) x US-09-816-494-3 (1-1998)  
QY 30 GlnLeuMetGlnGlnAnGlyIleGlyThrValLeuAsnAlaserAsnThrCysProLys 49  
DB 532 GAGCTGTATGAGCAGAGATGGATGTTATGTGTTAAATGCCAGCAATACCTGTCACAAAG 591  
QY 50 ProAspPheIleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGlu 69

Db 592 CCGTACTTATCCCGAGTCTCATTTCTCGGTGCTGTAATGACAGCTTTTGTGAG 651  
Qy 70 LysIleLeuProTrrLeuAspLysSerValAspPheIleGluLysAlaLysAlaSerAsn 89  
Db 652 AAAATTTTGGCGTGTGGACAAATCAGTAGATTTCATTGAGAAAGCAAAGCCTCCAA 711  
Qy 90 GlyCysValLeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAla 109  
Db 712 GGATGTGTTCTAGTGACACTGTTTCTAGCTGGATCTCCCGCTCCGCCACCATCGCTATGCC 771  
Qy 110 TyrIleMetLysArgMetAspMetSerLeuAspGluAlaTyrArgPheValLysGluLys 129  
Db 772 TACATCATGAGAGGATGACATGCTCTTTAGATGAGCTTACATGTTGTGAAGAAAAA 831  
Qy 130 ArgProThrIleSerProAsnPhenPhenLeuGlyGlnLeuLeuAspTyrGluLysLys 149  
Db 832 AGACCTACTATATCTCCAAACTTCAATTTCTGGGCCAACTCTCTGGACTATGAGAGAA 891  
Qy 150 IleLysAsnGlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuHisLeuGlu 169  
Db 892 ATTAAGACCAGACTGGAGCATCAGGGCCAAAGCAAACTCAAGCTGCTGCACCTGGAG 951  
Qy 170 LysProAsnGluProValProAlaValSerGluGlyGlnLysSerGluThrProLeu 189  
Db 952 AAGCCAAATGAACTGCTCCTCTCTCAGAGGCTGGACAGAAAGCAGAGCGCCCTC 1011  
Qy 190 SerProProCysAlaAspSerAlaThrSerGluAlaGlyGlnArgProValHisPro 209  
Db 1012 AGTCACCTCTGGCGATCTCTCTACCTCAGAGGAGCAGCAAAAGCCGCTGCACTCC 1071  
Qy 210 AlaSerValProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuVal 229  
Db 1072 GCCAGCGTGCCAGCGTGCCAGCGTGCCAGCGTCTGCTGTTAGAGGACAGCCCGCTG 1131  
Qy 230 GlnAlaLeuSerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnLysLeuLys 249  
Db 1132 CAGGCGCTCAGTGGCTGCACCTGTCACAGACAGCTGGAGACACAAATAGCTCAAG 1191  
Qy 250 ArgSerPheSerLeuAspIleLysSerValSerTyrSerAlaSerMetAlaAlaSerLeu 269  
Db 1192 CGTCTCTCTCTCGATATCAATCAGTTTCATATTCAGCCAGCATGCGAGCATCTTA 1251  
Qy 270 HisGlyPheSerSerGluAspAlaLeuGluTyrTyrLysProSerThrThrLeuAsp 289  
Db 1252 CATGCTTCTCTCATCAGAGATCTTTGGAATACTACAACTCTCCACTCTGAT 1311  
Qy 290 GlyThrAsnLysLeuCysGlnPheSerProValGlnGluLeuSerGluGlnThrProGlu 309  
Db 1312 GGGACCAACAGCTATGCGAGTTCTCCCTGTTTCAAGNACTATCGAGCAGACTCCGAA 1371  
Qy 310 ThrSerProAspLysGluGluAlaSerIleProLysLysLeuGlnThrAlaArgProSer 329  
Db 1372 ACCAGTCTGTATAGGAGGAGCAGCATCCCAAGAGCTGCAGACCGCCAGGCTTCA 1431  
Qy 330 AspSerGlnSerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArg 349  
Db 1432 GACAGCCAGCAAGCATGTCATTCGTTCAGACCCAGCAGCAGTGGCCGCCAGAGG 1491  
Qy 350 SerLeuLeuSerProLeuHisArgSerGlySerValGluAspAsnTyrHisThrSerPhe 369  
Db 1492 TCCCTTTATCTCCACTGCATCGAGTGGAGCGTGGAGGACAATACCAACAGCTTC 1551  
Qy 370 LeuPheGlyLeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLys 389  
Db 1552 CTTTTCGGCTTCTCCAGCAGCAGCAGCCTCAGAGTCTGCTGGCTTGGGCTTAA 1611  
Qy 390 GlyTyrHisSerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTyr 409  
Db 1612 GGTGGCACTCGGATATCTTGGCCCCCAGACCTTACCTCTCCCTGACCAAGAGCTGG 1671  
Qy 410 TyrPheAlaThrGluSerSerHisPheTyrSerAlaSerAlaIleTyrGlySerAla 429  
Db 1672 TATTTTGGCCAGAGTCTCTACACTTCTACTCTGCTCAGCCATCTACGGAGGAGTGCC 1731

RESULT 3

US-10-377-072-27  
; Sequence 27, Application US/10377072  
; Publication No. US20040009501A1  
; GENERAL INFORMATION:  
; APPLICANT: Millennium Pharmaceuticals Inc.  
; APPLICANT: Curtis, Rory A.J.  
; APPLICANT: Logan, Thomas Joseph  
; APPLICANT: Glucksmann, Maria A.  
; APPLICANT: Meyers, Rachel E.  
; APPLICANT: Williamson, Mark J.  
; APPLICANT: Rudolph-Owen, Laura A.  
; APPLICANT: Chun, Miyoung  
; APPLICANT: Tsai, Pong-Ying  
; TITLE OF INVENTION: NOVEL 25869, 25934, 26335, 50365, 21117,  
; TITLE OF INVENTION: 38692, 46508, 16816, 16839, 49937, 49931 AND 49933 MOLECULES  
; TITLE OF INVENTION: AND USES THEREFOR  
; FILE REFERENCE: MPI03-0180NM  
; CURRENT APPLICATION NUMBER: US/10/377,072  
; CURRENT FILING DATE: 2003-02-27  
; PRIOR APPLICATION NUMBER: US 09/895,860  
; PRIOR FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: US 60/215,370  
; PRIOR FILING DATE: 2000-06-29  
; PRIOR APPLICATION NUMBER: US 09/723,806  
; PRIOR FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: US 60/187,455  
; PRIOR FILING DATE: 2000-03-07  
; PRIOR APPLICATION NUMBER: US 09/843,297  
; PRIOR FILING DATE: 2001-04-25  
; PRIOR APPLICATION NUMBER: US 60/199,801  
; PRIOR FILING DATE: 2000-04-26  
; PRIOR APPLICATION NUMBER: US 09/861,801  
; PRIOR FILING DATE: 2001-05-21  
; PRIOR APPLICATION NUMBER: US 60/205,508  
; PRIOR FILING DATE: 2000-05-19  
; PRIOR APPLICATION NUMBER: US 09/816,494  
; PRIOR FILING DATE: 2001-03-23  
; PRIOR APPLICATION NUMBER: US 09/815,419  
; PRIOR FILING DATE: 2001-03-22  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 114  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 27  
; LENGTH: 1998  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(1998)  
US-10-377-072-27

## Alignment Scores:

```
Pred. No.: 0 Length: 1998
Score: 488.00 Matches: 488
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 94.39% Indels: 0
DB: 15 Gaps: 0

US-09-964-277-21 (1-517) x US-10-377-072-27 (1-1998)

Qy 30 GluLeuMetGlnGlnAenGlyIleGlyTyrValLeuAsnAlaSerAsnThrCysProlys 49
Db 532 GAGCTGATGACAGAGATGGATGGTATGTGTAAATGCCAGCAATACCTGTCCAAAG 591
Qy 50 ProAspPheIleProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlu 69
Db 592 CCGTGAATTTATCCCGAGTCTCATTTCCCTGCGTGTGCTGTAATGACAGCTTTTGTGAG 651
Qy 70 LysIleLeuProTrpLeuAspLysSerValAspPheIleGluLysAlaLysAlaSerAsn 89
Db 652 AAAATTTTGGCGTGGTGGCAAAATCAGTAGATTTCTTTGAGAAAGCAAAAGCCTCCAAAT 711
Qy 90 GlyCysValLeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAla 109
Db 712 GGATGTGTTCTAGTGCACTGTTTAGCTGGGATCTCCGCTCCGCCACCATCGCTATCGCC 771
Qy 110 TyrIleMetLysArgMetAspMetSerLeuAspGluAlaTyrArgPheValLysGluLys 129
Db 772 TAGCATCATGAAGAGATGGACATGCTTTTAGATGAAGCTTACAGATTTGTGAAAGAAAAA 831
Qy 130 ArgProThrIleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGluLysLys 149
Db 832 AGACCTACTATATCTCAAACTTCAATTTCTGGGCCAACTCCTGGACTATGAGAAGAAG 891
Qy 150 IleLysGlnGlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuHisLeuGlu 169
Db 892 ATTAAGAACCAGACTGGAGCATCAGGGCCAAAGAGCAAACTCAAGCTGTCACCTGGAG 951
Qy 170 LysProAsnGluProValProAlaValSerGluGlyGlyGlnLysSerGluThrProLeu 189
Db 952 AGCCAAATGAACTGCTCCCTGCTGTCTCAGAGGTGGACAGAAAGGAGAGCGCCCTC 1011
Qy 190 SerProCysAlaAspSerAlaThrSerGluAlaAlaGlyGlnArgProValHisPro 209
Db 1012 AGTCCACCCCTGTGCCACTCTGCTACTCAGAGGAGCAGAGCAAAAGGCCCGCTGCATCCC 1071
Qy 210 AlaSerValProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuVal 229
Db 1072 GCCAGCGTCCCGAGGTGCCACGCTGCGAGCGCTGCTGTTAGAGGACAGCCCGCTGATA 1131
Qy 230 GlnAlaLeuSerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnLysLeuLys 249
Db 1132 CAGCGCTCAGTGGGCTGCACCTGTCCGACAGACAGGCTGGAAGACAGCAATAAGCTCAAG 1191
Qy 250 ArgSerPheSerIleuAspLysSerValSerValSerAlaSerMetAlaAlaSerLeu 269
Db 1192 CGTTCCTCTCTCTGGGATATCAATCAGTTTCATATTCAGCCAGCATGGCAGCATCCTTA 1251
Qy 270 HisGlyPheSerSerGluAspAlaLeuGluTyrTyrLysProSerThrThrLeuAsp 289
Db 1252 CATGGCTTCTCCCTCATCAGAGATGCTTTGGATATCTACAACTTCCACTACTCTGGAT 1311
Qy 290 GlyThrAsnLysLeuCysGlnPheSerProValGlnGluLeuSerGluGlnThrProGlu 309
Db 1312 GGGACCAACAGCTATGCCAGTTCTCCCTGTTCAGGAACTATCGGAGCAGACTCCCGAA 1371
Qy 310 ThrSerProAspLysGluAlaSerIleProLysLysLeuGlnThrAlaArgProSer 329
Db 1372 ACCAGTCTCTATAGGAGGAGGAGCAGCATCCCGAAGACTGCAGACCCCGAGGCTTCA 1431
Qy 330 AspSerGlnSerLysArgLeuHisSerValArgThrSerSerGlyThrAlaGlnArg 349
Db 1432 GACAGCCAGAGCAAGCGATTGCAATTCGGTCAAGACCAAGCAGCAGTGGCACCCCGCAGAGG 1491
```

```
Qy 350 SerLeuLeuSerProLeuHisArgSerGlySerValGluAspAsnTyrHisThrSerPhe 369
Db 1492 TCCCTTTTATCTCCACTGCATCGAAGTGGGAGCGTGGAGGACAAATTACCAACAGCTTC 1551
Qy 370 LeuPheGlyLeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLeuLys 389
Db 1552 CTTTTCGGCCTTTCCACACGACGACGACCTCAGAACTCTGTGGCTGGGCTTAAG 1611
Qy 390 GlyTrpHisSerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTrp 409
Db 1612 GGTGGCACTCGGATATCTTGCCGCCCCAGACCTCTACCCCTTCCCTGACCAAGCTGG 1671
Qy 410 TyrPheAlaThrGluSerSerHisPheTyrSerAlaSerAlaIleTyrGlySerAla 429
Db 1672 TATTTTGGCAGAGAGTCTCTCACCTTCTACTCTGCTCAGCACTCTACGAGGAGGAGTGC 1731
Qy 430 SerTyrSerAlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerVal 449
Db 1732 AGTTACTCTGCTACAGCTGCAGCCAGCTGCCACTTGGCGAGACCAAGTCTATTCTGTG 1791
Qy 450 ArgArgArgGlnLysProSerAspArgAlaAspSerArgArgSerTrpHisGluGluSer 469
Db 1792 CGCAGCGCGCAAGCCAAAGTACAGAGCTCAGTCGCGGCGAGCTGGCATGAAGAGAGC 1851
Qy 470 ProPheGluLysGlnPheLysArgArgSerCysGlnMetGluPheGlyGluSerIleMet 489
Db 1852 CCTTTGAAAGACAGTTTAAAGCCAGAGCTGCCAATGGAAATTTGGAGAGCATCATG 1911
Qy 490 SerGluAsnArgSerArgGluGluLeuGlyLysValGlySerGlnSerSerPheSerGly 509
Db 1912 TCAGAGAACAGCTCAGCGGAGAGCTGGGGAAGTGGGCGAGTCACTGTCTAGCTTTTCGGGC 1971
Qy 510 SerMetGluIleIleLeuValSer 517
Db 1972 AGCATGGAATCATTTAGGTCTCC 1995
```

## RESULT 4

```
US-10-094-749-673
; Sequence 673, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAL, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NACHIKO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOKYU
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 673
; LENGTH: 2102
; TYPE: DNA
```



ORGANISM: Homo sapiens  
US-10-094-749-673

## Alignment Scores:

Pred. No.: 0 Length: 2102  
Score: 488.00 Matches: 488  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 94.39% Indels: 0  
DB: 15 Gaps: 0

US-09-964-277-21 (1-517) x US-10-094-749-673 (1-2102)

```
QY 30 GluLeuMetGlnGlnAsnGlyLeuGlyTyrValLeuLeuAlaSerAsnThrCysProLys 49
DB 587 GAGCTGATGACAGCAAGATGGGATTTGTTATGTTAAATGCCAGCAATACCTGTCACAAAG 646
QY 50 ProAspPheLeuProGluSerHisPheLeuArgValProValLeuAspSerPheCysGlu 69
DB 647 COTGACTTTATCCCGAGTCTCATTTCTCGCTGTGCTGTGAATGACAGCTTTTGTGAG 706
QY 70 LysIleLeuProTyrLeuAspLysValAspPheIleGluLysAlaLysAlaSerAsn 89
DB 707 AAAATTTTCCGCTGTGGCAATACAGTAGATTTCAATTGAGAAAGCAAAAGCTTCAT 766
QY 90 GlyCysValLeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAla 109
DB 767 GGATGTGTTAGTGCACCTGTTTGTAGCTGGGATCTCCGCTCCGCCACCATCGCTATCGCC 826
QY 110 TyrIleMetLysArgMetAspMetSerLeuAspGluAlaTyrArgPheValLysGluLys 129
DB 827 TACATCATGAAGAGATGGACATGCTTTTAGATGAAGCTTACAGATTGTGAAGAAAGAA 886
QY 130 ArgProThrIleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGluLysLys 149
DB 887 AGACCTACTATATCTCCAACTTCAATTTCTGGCCACTCTCTGGACTATGAGAAAG 946
QY 150 IleLysAsnGlnThrGlyAlaSerGlyProLysSerLysIleLysLeuLeuHisLeuGlu 169
DB 947 APTAAGAACCAAGCTGGAGCATCAGGGCAAGAGCAAAAGCAAACTCAAGCTGTCACCTGGAG 1006
QY 170 LysProAsnGluProValProAlaValSerGluGlyGlnLysSerGluThrProLeu 189
DB 1007 AAGCCAAATGAACCTGCTCTGCTGTCTCAGAGGTGACAGAAAGCGAGAGCGCCCTC 1066
QY 190 SerProCysAlaAspSerAlaThrSerGluAlaGlyGlnArgProValHisPro 209
DB 1067 AGTCCACCTGTGCCGACTCTGCTACCTCAGAGCGACAGCAAAAGGCCCGCTGCATCC 1126
QY 210 AlaSerValProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuVal 229
DB 1127 GCCAGCGTCCCGAGCGTCCCGAGCGTGCAGCGCTGCTGTTTAGAGGACAGCCCGCTGGTA 1186
QY 230 GlnAlaLeuSerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnLysLeuLys 249
DB 1187 CAGGCGCTCAGTGGCTGCACCTGTCGACAGAGGTGGAAGACAGCAATAAGCTCAAG 1246
QY 250 ArgSerPheSerLeuAspIleLysSerValSerTyrSerAlaSerMetAlaAlaSerLeu 269
DB 1247 CGTTCCTCTCTCTGGATCAAAATCAGTTTCATATTACGCCAGCATGGCAGCATCCTTA 1306
QY 270 HisGlyPheSerSerGluAspAlaLeuGluTyrTyrLysProSerThrThrLeuAsp 289
DB 1307 CATGGCTTCTCTCATCAGAGATGCTTTGGAATACTACAACCTTCATCTATCTGGAT 1366
QY 290 GlyThrAsnLysLeuCysGlnPheSerProValGlnGluLeuSerGluGlnThrProGlu 309
DB 1367 GGGACCAACAGCATATGCCAGTCTTCCCTGTTCCAGGAATATCGGAGCAGACTCCCGAA 1426
QY 310 ThrSerProAsnLysGluAlaSerIleProLysLysLeuGlnThrAlaArgProSer 329
DB 1427 ACCAGTCTGTATGAGAGAGCCAGCATCCCAAGAGCTGCCAGACTGCCAGCTGCCAGCTTCA 1486
```

```
QY 330 AspSerGlnSerLysArgLeuHisSerValArgThrSerSerGlyThrAlaGlnArg 349
DB 1487 GACAGCCAGAGCAGCATTCGATTCGTCAGAAACCAAGCAGCAGTGGCACCAGAGG 1546
QY 350 SerLeuLeuSerProLeuHisArgSerGlySerValGluAspAsnTyrHisThrSerPhe 369
DB 1547 TCCCTTTTATCTCCATGTCATCGAAGTGGAGCGGTGGAGACAATTACCACACAGCTTC 1606
QY 370 LeuPheGlyLeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLeuLys 389
DB 1607 CTTTTCGGCTTTTCCACGACGACGACCTCACGAAGTCTGCTGGCTGGGCCTTAAG 1666
QY 390 GlyThrHisSerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTrp 409
DB 1667 GCTGGCACTCGGATATCTTGGCCCCCAGACCTCTACCCCTTCCCTGACGACAGCTGG 1726
QY 410 TyrPheAlaThrGluSerSerHisPheTyrSerAlaSerAlaIleTyrGlyGlySerAla 429
DB 1727 TATTTGCCACAGAGTCTCTCACACTTCTACTCTGCTCAGCCATCTACGAGGACGTGCC 1786
QY 430 SerTyrSerAlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerVal 449
DB 1787 AGTTACTCTGCTACAGCTGCAGCCAGCTGCCACTTGGGAGACCAAGTCTATTCTGTG 1846
QY 450 ArgArgArgGlnLysProSerAspArgAlaAspSerArgArgSerTrpHisGluGluSer 469
DB 1847 CGCAGCGCGCAGAACCAAGTGCAGAGCTGACTCGCGCGGAGCTGGCATGAAGAGAGC 1906
QY 470 ProPheGluLysGlnPheLysArgArgSerCysGlnMetGluPheGlyGluSerIleMet 489
DB 1907 CCTTTGAAAGCAGTTTAAACCGCAGAGCTGCCAATGGAAATTTGGAGAGCATCATG 1966
QY 490 SerGluAsnArgSerArgGluGluLeuGlyLysValGlySerGlnSerSerPheSerGly 509
DB 1967 TCAGAGAACAGCTCAGCGGAGAGCTGGGAAAAGTGGGAGTGGCAGTCACTAGCTTTTCGGGC 2026
QY 510 SerMetGluIleIleGluValSer 517
DB 2027 AGCATGGAATCATTTAGGTCTCC 2050
```

## RESULT 5

US-09-964-277-1  
; Sequence 1, Application US/09964277  
; Patent No. US20020137170A1  
; GENERAL INFORMATION:  
; APPLICANT: Lucche, Ralf M.  
; APPLICANT: Wei, Bo  
; TITLE OF INVENTION: DSP-16 DUAL-SPECIFICITY PHOSPHATASE  
; FILE REFERENCE: 200125.434  
; CURRENT APPLICATION NUMBER: US/09/964,277  
; CURRENT FILING DATE: 2001-09-25  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 3496  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-964-277-1

Alignment Scores:  
Pred. No.: 0 Length: 3496  
Score: 488.00 Matches: 488  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 94.39% Indels: 0  
DB: 9 Gaps: 0

US-09-964-277-21 (1-517) x US-09-964-277-1 (1-3496)

```
QY 30 GluLeuMetGlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerAsnThrCysProLys 49
DB 1093 GAGCTGATGACAGCAAGATGGGATTTGTTATGTTAAATGCCAGCAATACCTGTCACAAAG 1152
```

QY 50 ProAspPheIleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGlu 69  
DB 1153 CTGACTTATCCCGAGTCATCTTCTGCGTGTGCGCTGTGAATGACAGCTTTGTGAG 1212  
QY 70 LysIleLeuProTrpLeuAspLysSerValAspPheIleGluLysAlaLysAlaSerAsn 89  
DB 1213 AAAATTTTCCCGTGTGTGACAAATCAGTAGATTTCATTGAGAAAGCAAGCTCCCAAT 1272  
QY 90 GlyCysValLeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAla 109  
DB 1273 GGATGTGTTCTAGTGCACCTGTTTACCTGGATCTCCGCTCCGCCACCATCGCTATCGCC 1332  
QY 110 TyrIleMetLysArgMetAspMetSerLeuAspGluAlaTyrArgPheValLysGluLys 129  
DB 1333 TACATCATGAAGAGATGACATGTCCTTTAGATGAAGCTTACAGATTGTGAAGAAAAA 1392  
QY 130 ArgProThrIleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGluLysLys 149  
DB 1393 AGACCTACTATATCTCCAACTTCATTTCTGGGCCAACTCCTGGACTATGAGAAGAAG 1452  
QY 150 IleLysAsnGlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuLeuHisLeuGlu 169  
DB 1453 ATTAGAACCAGATGGAGCATCAGGGCCAAAGAGCAAACTCAAGCTGCTGCACCTGGAG 1512  
QY 170 LysProAsnGluProValProAlaValSerGluGlyGlnLysSerGluThrProLeu 189  
DB 1513 AAGCCAAATGAACCTGCTCCTGCTGCTCAGAGGTGACAGAAAGAGGAGCGCCCTC 1572  
QY 190 SerProProCysAlaAspSerAlaThrSerGluAlaAlaGlyGlnArgProValHisPro 209  
DB 1573 AGTCCACCTGTGCGGACTCTGCTACCTCAGAGGCGAGGACAAAGCCCGTGCATCCC 1632  
QY 210 AlaSerValProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuVal 229  
DB 1633 GCAGCGTGCCTGAGCGTGCCTGAGCGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1692  
QY 230 GlnAlaLeuSerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnLysLeuLys 249  
DB 1693 CAGCGCTCAGTGGGCTGCACCTGCTCGCAGACAGCGTGGAGACAGCAATAAGCTCAAG 1752  
QY 250 ArgSerPheSerLeuAspLysSerValSerTyrSerAlaSerMetAlaAspSerLeu 269  
DB 1753 CGTTCTCTCTCTGATATCAAACTCAGTTTCATATTTCAGCCAGCATGGCAGCATCTTA 1812  
QY 270 HisGlyPheSerSerGluAspAlaLeuGluTyrTyrLysProSerThrThrLeuAsp 289  
DB 1813 CATGGCTTCTCCTCATCAGAGATGCTTTGGATATCTACAAACCTTCCACTACTCGAT 1872  
QY 290 GlyThrAsnLysLeuCysGlnPheSerProValGlnGluLeuSerGluGlnThrProGlu 309  
DB 1873 GGGACCAACAAGCTATGCCAGTTCTCCCTGTTTCAGGAATATCGGAGCAGACTCCCGAA 1932  
QY 310 ThrSerProAspLysGluGluAlaSerIleProLysLysLeuGlnThrAlaArgProSer 329  
DB 1933 ACCAGTCTCTGATGAAGAGGAGCCAGCATCCCAAGAGCTGCAGACCGCCAGGCTTCA 1992  
QY 330 AspSerGlnSerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArg 349  
DB 1993 GACAGCCAGAGCAAGCAATGCAATTCGCTCAGAACCCAGCAGCAGTGGCCGCCAGAGG 2052  
QY 350 SerLeuLeuSerProLeuHisArgSerGlySerValGluAspAsnTyrHisThrSerPhe 369  
DB 2053 TCCCTTTTATCTCCATGCAATCGAGTGGAGCGGAGGACAAATACACACAGCTTC 2112  
QY 370 LeuPheGlyLeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLys 389  
DB 2113 CTTTTGGCCCTTCCACCGCCAGCAGCACCTCAGAGTCTGCTGGCTGGCTGAGCTTAAG 2172  
QY 390 GlyTrpHisSerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTrp 409  
DB 2173 GGTGGCAGCTCGGATATCTTGGCCCCCAGACCTTACCTCCCTTCCCTGACGAGCTGG 2232  
QY 410 TyrPheAlaThrGluSerSerHisPheTyrSerAlaSerAlaIleTyrGlySerAla 429

DB 2233 TATTTTGCACAGAGTCTCTCACACTTCTACTCTGCTCAGCCATCTACGGAGCAGTGCC 2292  
QY 430 SerTyrSerAlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerVal 449  
DB 2293 AGTTACTCTGCTACAGCTGACCCAGCTGCCACATTCGGAGACCAAGTCTATTCTGTG 2352  
QY 450 ArgArgGlnLysProSerAspArgAlaAspSerArgSerTrpHisGluGluSer 469  
DB 2353 CGCAGGCGCAGAAAGCCAAAGTCAGAGCTGACTCGCGCGGAGCTGGCATGAAGAGAGC 2412  
QY 470 ProPheGluLysGlnPheLysArgArgSerCysGlnMetGluPheGlyGluSerIleMet 489  
DB 2413 CCCTTTGAAGAGCAGTTTAAACGCAAGAGCTGCCAATGGAAATTTGGAGAGACATCATG 2472  
QY 490 SerGluAsnArgSerArgGluGluLeuGlyLysValGlySerGlnSerSerPheSerGly 509  
DB 2473 TCAGAGAACAGCTCACGGAAGAGCTGGGAAAGTGGCAGTCAGTCTAGCTTTTCGGGC 2532  
QY 510 SerMetGluIleIleGluValSer 517  
DB 2533 AGCATGGAATCATTTAGGTCTCC 2556  
RESULT 6  
US-09-816-494-1  
; Sequence 1. Application US/09816494  
; Patent No. US20020034807A1  
; GENERAL INFORMATION:  
; APPLICANT: Meyers, Rachel A.  
; TITLE OF INVENTION: 38692 AND 21117, NOVEL DUAL SPECIFICITY  
; TITLE OF INVENTION: PHOSPHATASE MOLECULES AND USES THEREFOR  
; FILE REFERENCE: 10448-030002  
; CURRENT APPLICATION NUMBER: US/09/816,494  
; CURRENT FILING DATE: 2001-03-23  
; PRIOR APPLICATION NUMBER: US 60/191,858  
; PRIOR FILING DATE: 2000-03-24  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 3544  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (589)...(2583)  
US-09-816-494-1  
Alignment Scores:  
Pred. No.: 0 Length: 3544  
Score: 488.00 Matches: 488  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 94.39% Indels: 0  
DB: Gaps: 0  
US-09-964-277-21 (1-517) x US-09-816-494-1 (1-3544)  
QY 30 GluLeuMetGlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerAsnThrCysProLys 49  
DB 1120 GAGCTGATGCAGCAATGGGATTTGTTATGTGTTAAATGCCAGCAATACCTGTCCAAAG 1179  
QY 50 ProAspPheIleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGlu 69  
DB 1180 CTGACTTTATCCCGAGTCTCAATTTCTGCGTGTGCTGTGAATGACAGCTTTGTGAG 1239  
QY 70 LysIleLeuProTrpLeuAspLysSerValAspPheIleGluLysAlaLysAlaSerAsn 89  
DB 1240 AAAATTTTCCCGTGTGTGACAAATCAGTAGATTTCATTGAGAAAGCAAGCTCCCAAT 1299  
QY 90 GlyCysValLeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAla 109  
DB 1300 GGATGTGTTCTAGTGCATGTTTACTGGGATCTCCGCTCCGCCACCATCGCTATCGCC 1359

QY 110 TyrIleMetLysArgMetAspMetSerLeuAspGluAlaTyrArgPheValLysGluLys 129  
DB 1360 TACATCATGAAGAGATGGACATCTCTTTAGATGAAGCTTTACAGATTTGTGAAGAAAA 1419  
QY 130 ArgProThrIleSerProAspPheLeuGlyGlnLeuLeuAspTyrGluLysLys 149  
DB 1420 AGACCTACTATATCTCCAACTTCAATTTCTGGCCAACTCTCTGACTATGAGAGAG 1479  
QY 150 IleLysAsnGlnThrGlyAlaSerGlyProLysSerLysLysLeuLysLeuGlu 169  
DB 1480 ATTAAGAACCAAGATGGAGCATCAGGCGCAAGAGCAAACTCAAGCTGTGCACTGGAG 1539  
QY 170 LysProAsnGluProValProAlaValSerGluGlyGlnLysSerGluThrProLeu 189  
DB 1540 AAGCCAAATGAACCTGCTCTCTGCTGCTCAGAGGCTGCAGAAAGCGAGAGCCCTTC 1599  
QY 190 SerProProCysAlaAspSerAlaThrSerGluAlaGlyGlnArgProValHisPro 209  
DB 1600 AGTCCACCCCTGCGGACTCTGCTACCTCAGAGCGAGCAGCAAAAGCGCCGTGCATCC 1659  
QY 210 AlaSerValProSerValProSerValGlnProSerLeuLeuAspSerProLeuVal 229  
DB 1660 GCCAGCGTGGCCAGCGTCCAGCGTGCAGCGCTGCTGTAGAGGACACCCCGCTGGTA 1719  
QY 230 GlnAlaLeuSerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnLysLeuLys 249  
DB 1720 CAGCGCTCAGTGGCTGACCTGCTCCGAGACAGGCTGGAGACAGCAATAAGCTCAAG 1779  
QY 250 ArgSerPheSerLeuAspIleLysSerValSerTyrSerAlaSerMetAlaAlaSerLeu 269  
DB 1780 CGTCTCTCTCTCTGATATCAATATCAATATCAATATCAATATCAATATCAATATCA 1839  
QY 270 HisGlyPheSerSerGluAspAlaLeuGluTyrTyrLysProSerThrThrLeuAsp 289  
DB 1840 CATGCTCTCTCTCTATCAGAGAGTCTTGGAACTACTCAAACTTCCACTCTCTGGAT 1899  
QY 290 GlyThrAsnLysLeuCysGlnPheSerProValGlnGluLeuSerGluGlnThrProGlu 309  
DB 1900 GCGACCAACAGCTATGCGAGTCTCTCCCTGCTTCCAGGAATCTATCGGAGCAGACTCC 1959  
QY 310 ThrSerProAspLysGluGluAlaSerIleProLysLysLeuGlnThrAlaArgProSer 329  
DB 1960 ACCAGTCTGTATGAGAGAGAGCAGCATCCCAAGAGCTGCAGACCGCAGCGCTTCA 2019  
QY 330 AspSerGlnSerLysArgLeuHisSerValArgThrSerSerGlyThrAlaGlnArg 349  
DB 2020 GACACCCAGAGCAGCGATTCATTCGTCAGAACACAGCAGCAGTGCACCGCCAGAGG 2079  
QY 350 SerLeuLeuSerProLeuHisArgSerGlySerValGluAspAsnTyrHisThrSerPhe 369  
DB 2080 TCCCTTTTATCTCCACTGCTATCGAAGTGGAGCGTGGAGCAATATACCAACACAGCTTC 2139  
QY 370 LeuPheGlyLeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuLys 389  
DB 2140 CTTTTCGGCTTTCACACAGCAGCAGCACTCAGAGTCTGTCGCTGGCCCTTAAG 2199  
QY 390 GlyTrpHisSerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerTrp 409  
DB 2200 GCGTGGCACTCGATATCTTGGCCCCCAGACCTCTACCCCTTCCCTGACCAAGCGTGG 2259  
QY 410 TyrPheAlaThrGluSerSerHisPheTyrSerAlaSerAlaIleTyrGlyGlySerAla 429  
DB 2260 TATTTTGGCCAGAGTCTCTACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2319  
QY 430 SerTrpSerAlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerVal 449  
DB 2320 AGTTACTCTCTCTCAGCTGAGCCAGCTGCGCCTGCGGAGACCAAGCTATTCTGTG 2379  
QY 450 ArgArgArgGlnLysProSerAspArgAlaAspSerArgArgSerTrpHisGluGluSer 469  
DB 2380 CGCAGCGCGCAAGCGCAAGTACAGAGCTGACTCGCGCGGAGCTGCGATGAGAGAGC 2439  
QY 470 ProPheGluLysGlnPheLysArgArgSerCysGlnMetGluPheGlyGluSerIleMet 489

DB 2440 CCCTTTGAAAGCAGTTTAAACGACAGAGCTGCCAATGGAATTTGGAGAGCATCATG 2499  
QY 490 SerGluAsnArgSerArgGluGluLeuGlyLysValGlySerGlnSerSerPheSerGly 509  
DB 2500 TCAGAGAACAGGTCACGGGAAGAGCTGGGGAAGAGTGGGGAAGAGTGGGAGTCTAGCTTTTCGGGC 2559  
QY 510 SerMetGluIleIleGluValSer 517  
DB 2560 AGCATGGAATCATTTAGGTTCTCC 2583  
RESULT 7  
US-10-377-072-25  
; Sequence 25, Application US/10377072  
; Publication No. US20040009501A1  
; GENERAL INFORMATION:  
; APPLICANT: Millennium Pharmaceuticals Inc.  
; APPLICANT: Curtis, Rory A.J.  
; APPLICANT: Logan, Thomas Joseph  
; APPLICANT: Glucksmann, Maria A.  
; APPLICANT: Meyers, Rachel B.  
; APPLICANT: Williams, Mark J.  
; APPLICANT: Rudolph-Owen, Laura A.  
; APPLICANT: Chun, Miyoung  
; APPLICANT: Tsai, Fong-Ying  
; TITLE OF INVENTION: NOVEL 25869, 25934, 26335, 50365, 21117,  
; TITLE OF INVENTION: 38592, 46508, 16816, 16839, 49937, 49931 AND 49933 MOLECULES  
; TITLE OF INVENTION: AND USES THEREFOR  
; FILE REFERENCE: MP103-0180NMIM  
; CURRENT APPLICATION NUMBER: US/10/377,072  
; CURRENT FILING DATE: 2003-02-27  
; PRIOR APPLICATION NUMBER: US 09/895,860  
; PRIOR FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: US 60/215,370  
; PRIOR FILING DATE: 2000-06-29  
; PRIOR APPLICATION NUMBER: US 09/723,806  
; PRIOR FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: US 60/187,455  
; PRIOR FILING DATE: 2000-03-07  
; PRIOR APPLICATION NUMBER: US 09/843,297  
; PRIOR FILING DATE: 2001-04-25  
; PRIOR APPLICATION NUMBER: US 60/199,801  
; PRIOR FILING DATE: 2000-04-26  
; PRIOR APPLICATION NUMBER: US 09/861,801  
; PRIOR FILING DATE: 2001-05-21  
; PRIOR APPLICATION NUMBER: US 60/205,508  
; PRIOR FILING DATE: 2000-05-19  
; PRIOR APPLICATION NUMBER: US 09/816,494  
; PRIOR FILING DATE: 2001-03-23  
; PRIOR APPLICATION NUMBER: US 09/815,419  
; PRIOR FILING DATE: 2001-03-22  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 114  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 25  
; LENGTH: 3544  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (589) ... (2586)  
US-10-377-072-25  
Alignment Scores:  
Pred. No.: 0  
Score: 488.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 94.39%  
DB: 15  
Length: 3544  
Matches: 488  
Conservative: 0  
Mismatch: 0  
Indels: 0  
Gaps: 0  
US-09-964-277-21 (1-517) x US-10-377-072-25 (1-3544)

Qy 30 GluLeuMetGlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerAsnThrCysProLys 49  
Db 1120 GAGCTCATGCAGCAGATGGGATGGTTATGTTGTTAAATGCAGCAATACCTGTCCAAAG 1179  
Qy 50 ProAspPheileProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlu 69  
Db 1180 CTGACTTATCCCGAGGCTCATTTCCGTGCGTGGCTGTGAATGACAGCTTTTGTGAG 1239  
Qy 70 LysIleLeuProTrpLeuAspLysSerValAspPheIleGluLysAlaLysAlaSerAsn 89  
Db 1240 AAAATTTTCCGTTGGTGGACAAATCAGTAGATTTCATTGAGAAAGCAAAAGCTCCAAAT 1299  
Qy 90 GlyCysValLeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAla 109  
Db 1300 GGATGTGTTCTAGTGACATGTTTAGCTGGGATCTCCCGCTCGCCACCATCGCTATCGCC 1359  
Qy 110 TyrIleMetLysArgMetAspMetSerLeuAspGluAlaTyrArgPheValLysGluLys 129  
Db 1360 TACATCATGAGAGGATGACATGCTTTAGATGAAGCTTACAGATTGTGAAGAAAAA 1419  
Qy 130 ArgProThrIleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGluLysIys 149  
Db 1420 AGACCTACTATATCTCAAACTTCAATTTCTGGCCAACTCTCGACTATGAGAAGAAG 1479  
Qy 150 IleLysAsnGlnThrGlyAlaSerGlyProLysSerLysLysLeuLysLeuLeuHisLeuGlu 169  
Db 1480 ATTAGAACCAGACTGGAGCATCAGGGCCAAAGACAACTCAAGTCTGCACTGGAG 1539  
Qy 170 LysProAsnGluProValProAlaValSerGluGlyGlnLysSerGluThrProLeu 189  
Db 1540 AAGCCAAATGAACCTGCTGCTGCTCAGAGGGTGGACAGAAAGGAGAGACGCCCTC 1599  
Qy 190 SerProCysAlaAspSerAlaThrSerGluAlaGlyGlnArgProValHisPro 209  
Db 1600 AGTCCACCTGTGCCACTCTGCTACCTCAGAGGAGGAGGACAAAGCCCGTGCATCCC 1659  
Qy 210 AlaSerValProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuVal 229  
Db 1660 GCCAGGGTGGCCAGCTGCCAGCTGACGGCTGCTGTATAGAGACAGCCCGCTGGTA 1719  
Qy 230 GlnAlaLeuSerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnLysLeuLys 249  
Db 1720 CAGCGCTCAGTGGGTGACCTGTCCGAGACAGGCTGGAAAGACAGCAATPAAGCTCAAG 1779  
Qy 250 ArgSerPheSerLeuAspLysSerValSerTyrSerAlaSerMetAlaAlaSerLeu 269  
Db 1780 CGTTCCTTCTCTGGATATCAATCAGTTTCATATTCAGCCAGCATGGCAGATCCTTA 1839  
Qy 270 HisGlyPheSerSerGluAspAlaLeuGluTyrTyrLysProSerThrThrLeuAsp 289  
Db 1840 CATGGCTTCTCCTCATCAGAAGATGTTTGGAAATACCTACAAACCTTCCACTACTCTGGAT 1899  
Qy 290 GlyThrAsnLysLeuCysGlnPheSerProValGlnGlnLeuSerGluGlnThrProGlu 309  
Db 1900 GGGNCCACAGCATATGCCAGTTCTCCCTGTTCCAGAACTATCGAGACAGACTCCCGAA 1959  
Qy 310 ThrSerProAspLysGluGluAlaSerIleProLysLysLeuGlnThrAlaArgProSer 329  
Db 1960 ACCAGTCTCTAATAGGAGGAAGCCAGCATCTCCCAAGAAGCTGCAGACCCGCCAGGCTTCA 2019  
Qy 330 AspSerGlnSerLysArgLeuHisSerValArgThrSerSerGlyThrAlaGlnArg 349  
Db 2020 GACAGCCAGACAGCGATTGCAATTGGTTCAGNACCAGCAGCAGTGGCCNCCGCCAGAGG 2079  
Qy 350 SerLeuLeuSerProLeuHisArgSerGlySerValGluAspAsnTyrHisThrSerPhe 369  
Db 2080 TCCCTTTTATCTCCACTGCATCGAAGTGGAGCGTGGAGGACAAATTACCACACAGCTTC 2139  
Qy 370 LeuPheGlyLeuSerThrSerGlnHisLeuThrLysSerAlaGlyLeuGluLysLys 389  
Db 2140 CTTTTCGGCTTTCCACAGCCAGCAGACCTCAGGAAGTCTGCTGGCTGGCCCTTAAG 2199  
Qy 390 GlyTrpHisSerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTrp 409

Db 2200 GGCTGGCACTCGGATATCTTGGCCCCCAGACCTCTACCCCTTCCCTGACAGAGCTGG 2259  
Qy 410 TyrPheAlaThrGluSerSerHisPheTyrSerAlaSerAlaIleTyrGlyGlySerAla 429  
Db 2260 TATTTTGGCAGAGAGTCTCTACACTTCTACTCTGCCTCAGCCATCTACGGAGGAGTCC 2319  
Qy 430 SerTyrSerAlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerVal 449  
Db 2320 AGTTACTCTGCTACAGCTGACAGCAGCTGCCACTTGGCGAGACCAAGTCTATTCTGTG 2379  
Qy 450 ArgAspArgGlnLysProSerAspArgAlaAspSerArgArgSerTrpHisGluGluSer 469  
Db 2380 CCACGCGCGAGAGCCAGTGCAGAGCTGACTCGCGCGGAGCTGCATGATGAGAGAGC 2439  
Qy 470 PropheGluLysGlnPheLysArgSerCysGlnMetGluPheGlyGluSerIleMet 489  
Db 2440 CCTTTTAAAAAGCAGTTTAAACGCGAGAGCTGCCAAATGGAAATTTGGAGAGAGCATCATG 2499  
Qy 490 SerGluAsnArgSerArgGluGluLeuGlyLysValGlySerGlnSerSerPheSerGly 509  
Db 2500 TCAGAGAACAGGTACCGGAAGAGCTGGGAAAGTGGCAGTCACTCTAGCTTTTCGGGC 2559  
Qy 510 SerMetGluIleIleGluValSer 517  
Db 2560 AGCATGGAAATCAATTGAGGTCTCC 2583  
RESULT 8  
US-10-425-114-26234  
; Sequence 26234, Application US/10425114  
; Publication NO. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 26234  
; LENGTH: 3625  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB4119-028-H6\_FLI  
US-10-425-114-26234  
Alignment Scores:  
Pred. No.: 0 Length: 3625  
Score: 488.00 Matches: 488  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 94.39% Indels: 0  
DB: 12 Gaps: 0  
US-09-964-277-21 (1-517) x US-10-425-114-26234 (1-3625)  
Qy 30 GluLeuMetGlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerAsnThrCysProLys 49  
Db 1223 GAGCTGATGCAGACAGATGGGATTTGGTTATGTGTTAATGCGAGCAATACCTGTCCAAAG 1282  
Qy 50 ProAspPheileProGlnSerHisPheLeuArgValProValAsnAspSerPheCysGlu 69  
Db 1283 CTTGACTTTATCCCGAGGCTCATTTCTGCTGGCTGTGGAATGACAGCTTTTGTGAG 1342  
Qy 70 LysIleLeuProTrpLeuAspLysSerValAspPheIleGluLysAlaLysAlaSerAsn 89  
Db 1343 AAAATTTTCCGTTGGTGGACAAATCAGTAGATTTCATTGAGAAAGCAAAAGCTCCAAAT 1402

QY 90 GlyCysValLeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAla 109  
Db 1403 GGATGTGTTCTAGTGCACTGTTTAGCTGGGATCTCCCGCTCCGCCACCATCGCTATCGCC 1462  
QY 110 TyrIleMetLysArgMetAspMetSerLeuAspGluAlaTyrArgPheValLysGluLys 129  
Db 1463 TACATCATGAAGAGATGGACATGTCCTTAGATGAAGCTTACAGATTGTGAAGAAGAAA 1522  
QY 130 ArgProThrIleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGluLysLys 149  
Db 1523 AGACCTACTATATCTCCAACTTCAATTTCTGGCCAACTCTCTGGACTATGAGAAGAAG 1582  
QY 150 IleLysAsnGlnThrGlyAlaSerGlyProLysSerLysLysLeuLysLeuHisLeuGlu 169  
Db 1583 ATTAAGAACCAGACTGGAGCATCAGGCCAAAGAGCAAACTCAAGCTGTCACCTGGAG 1642  
QY 170 LysProAsnGluProValProAlaValSerGluGlyGlnLysSerGluThrProLeu 189  
Db 1643 AAGCCAAATGAACCTGTCTCTGCTCTCAGAGGCTGCAGAAAAGCGAGACGCCCTC 1702  
QY 190 SerProProCysAlaAspSerAlaThrSerGluAlaGlyGlnArgProValHisPro 209  
Db 1703 AGTCCACCTGTGGCGACTCTCTGCTACCTCAGAGGAGCAGCAAGAGCCGCTGCATCC 1762  
QY 210 AlaSerValProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuVal 229  
Db 1763 GCCAGGTGCCAGGTCGCCAGGTGCCAGGCTGCCGCTGTAGAGGACGCCCGCTGGTA 1822  
QY 230 GlnAlaLeuSerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnLysLeuLys 249  
Db 1823 CAGGCGCTCAGTGGGCTGCACCTGTCCGACAGAGGCTGGAAGACAGCAATAAGCTCAAG 1882  
QY 250 ArgSerPheSerLeuAspIleLysSerValSerTyrSerAlaSerMetAlaLeuSerLeu 269  
Db 1883 CGTTCCTCTCTGGATCAATCATCATCTGTTCAATTACCGCAGCATGGCAGCATCTTA 1942  
QY 270 HisGlyPheSerSerSerGluAspAlaLeuGluTyrTyrLysProSerThrThrLeuAsp 289  
Db 1943 CATGCTTCTCTCATCAGAAGATGCTTTGGAACTACTACAACTCCACTACTCTGGAT 2002  
QY 290 GlyThrAsnLysLeuCysGlnPheSerProValGlnGluLeuSerGluGlnThrProGlu 309  
Db 2003 GGGACCAACAGCTATGCCAGTCTCCCTCTGTTCCAGAACTATCGGAGCAGACTCCCGAA 2062  
QY 310 ThrSerProAspLysGluGluAlaSerIleProLysLysLeuGlnThrAlaAspProSer 329  
Db 2063 ACCAGTCTGTATAGGAGAACCCAGCATCCCAAGAGCTGCAGACCGCCAGCCCTTCA 2122  
QY 330 AspSerGlnSerLysArgLeuHisSerValArgThrSerSerGlyThrAlaGlnArg 349  
Db 2123 GACAGCCAGAGCAAGCGATTGCAATTCGGTCAGAAACCAGCAGCAGTGCGCACCGCCAGAG 2182  
QY 350 SerLeuLeuSerProLeuHisArgSerGlySerValGluAspAsnTyrHisThrSerPhe 369  
Db 2183 TCCCTTTTATCTCCACTGCATCAGTGGAGTGCGAGTGAGGACAAATTACCAACACGCTTC 2242  
QY 370 LeuPheGlyLeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLeuLys 389  
Db 2243 CTTTTCGGCTTTCACCCAGCAGCAGCAGCATCCAGAGCTGCTGCTGCGCTGCGCTTAAG 2302  
QY 390 GlyThrHisSerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTrp 409  
Db 2303 GGCTGGCATCGGATATCTTGCCCCCAGACCTCTACCCCTTCCCTGACACGACGCTGG 2362  
QY 410 TyrPheAlaThrGluSerSerHisPheTyrSerAlaSerAlaIleTyrGlyGlySerAla 429  
Db 2363 TATTTTGGCAGAGTCTCATCTTCTACTCTGCTCAGCCATCTACGAGGAGCAGTGCC 2422  
QY 430 SerTyrSerAlaTyrSerSerGlnLeuProThrCysGlyAspGlnValTyrSerVal 449  
Db 2423 AGTTACTTGCTCAGCTCAGCGCAGCTGCGCCACTTGCAGCAACCAAGTCTATTCTGTG 2482

QY 450 ArgArgArgGlnLysProSerAspArgAlaAspSerArgArgSerTrpHisGluGluSer 469  
Db 2483 CGCAGGGCGCAGAGCAAGCAAGTGCAGAGCTGACTCGCGCGAGCTGGCATGTAAGAGAGC 2542  
QY 470 ProPheGluLysGlnPheLysArgArgSerCysGlnMetGluPheGlyGluSerIleMet 489  
Db 2543 CCCTTTGAAGAGCAGTTTAAACCGCAGAGCTGCCAATGGAAATTTGGAGAGACATCATG 2602  
QY 490 SerGluAsnArgSerArgGluGluLeuGlyLysValGlySerGlnSerSerPheSerGly 509  
Db 2603 TCAGAGAACAGCTCACCGGAAGAGCTGGGGAAGTGGGAAAAGTGGCAGTCAGTCTAGCTTTTCGGGC 2662  
QY 510 SerMetGluIleIleGluValSer 517  
Db 2663 AGCATGGAATCATTTGAGGTCTCC 2686  
RESULT 9  
US-10-072-012-255  
; Sequence 255, Application US/10072012  
; Publication No. US20040033493A1  
; GENERAL INFORMATION:  
; APPLICANT: Tchernev, Velizar  
; APPLICANT: Spytek, Kimberly  
; APPLICANT: Zerhusen, Bryan  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Shimkets, Richard  
; APPLICANT: Li, Li  
; APPLICANT: Gangoli, Esha  
; APPLICANT: Padigaru, Muralidhara  
; APPLICANT: Anderson, David W.  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Miller, Charles E.  
; APPLICANT: Gerlach, Valerie  
; APPLICANT: Taupier, Jr. Raymond J.  
; APPLICANT: Gusev, Vladimir Y.  
; APPLICANT: Colman, Steven D.  
; APPLICANT: Wolenc, Adam R.  
; APPLICANT: Pena, Carol E. A  
; APPLICANT: Furtak, Katarzyna  
; APPLICANT: Grosse, William M.  
; APPLICANT: Alsbrook II, John P.  
; APPLICANT: Lepley, Denise M.  
; APPLICANT: Rieger, Daniel K.  
; APPLICANT: Burgess, Catherine E.  
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same  
; FILE REFERENCE: 21402-258  
; CURRENT APPLICATION NUMBER: US/10/072,012  
; CURRENT FILING DATE: 2002-01-31  
; PRIOR APPLICATION NUMBER: 60/265,102  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: 60/265,514  
; PRIOR FILING DATE: 2001-01-31  
; PRIOR APPLICATION NUMBER: 60/265,517  
; PRIOR FILING DATE: 2001-01-31  
; PRIOR APPLICATION NUMBER: 60/265,412  
; PRIOR FILING DATE: 2001-01-31  
; PRIOR APPLICATION NUMBER: 60/265,395  
; PRIOR FILING DATE: 2001-01-31  
; PRIOR APPLICATION NUMBER: 60/266,406  
; PRIOR FILING DATE: 2001-02-02  
; PRIOR APPLICATION NUMBER: 60/266,767  
; PRIOR FILING DATE: 2001-02-05  
; PRIOR APPLICATION NUMBER: 60/267,057  
; PRIOR FILING DATE: 2001-02-07  
; PRIOR APPLICATION NUMBER: 60/266,975  
; PRIOR FILING DATE: 2001-02-07  
; PRIOR APPLICATION NUMBER: 60/267,459  
; PRIOR FILING DATE: 2001-02-08  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 1391  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 255  
; LENGTH: 2200

TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-072-012-255

Alignment Scores:  
Pred. No.: 0 Length: 2200  
Score: 408.00 Matches: 435  
Percent Similarity: 99.32% Conservative: 0  
Best Local Similarity: 99.32% Mismatches: 0  
Query Match: 78.92% Indels: 3  
DB: 12 Gaps: 0

US-09-964-277-21 (1-517) x US-10-072-012-255 (1-2200)

```
QY 83 GLuLYeAlaLYsAlaSerAenGLyCysValLeuValHisCysLeuAlaGLyILeSerArg 102
DB 793 GAGAAAGCAAAAGCCCTCAATGGATGTCTAGTGCACCTGTTAGCTGGGATCTCCGCG 852
QY 103 SerAlaThrIleAlaIleAlaTyrlleMetLYsArgMetAspMetSerLeuAspGluAla 122
DB 853 TCCGCCACCATCGCTATCGCTACATCATGAAGAGGATGGACATGCTTTTAGATGAAGCT 912
QY 123 Tyr---ArgPheValLYsGLuLYsArgProThrIleSerProAsnPheAsnPheLeuGLy 141
DB 913 TACAGGAGATTGTGAAGAAAAGACCTACTATATCTCCAAACTTCAATTTCTGGCG 972
QY 142 GLnLeuLeuAspTYrGLuLYsLYsIleLYsAsnGLnThrGLyAlaSerGLyProLYsSer 161
DB 973 CAACCTCTGGACTATGAGAAGAGAGATTAAAGACCAAGACTGGAGCATCAGGCCCCAAGAGC 1032
QY 162 LYsLeuLYsLeuLeuHisLeuGLuLYsProAsnGLuProValProAlaValSerGLuGLy 181
DB 1033 AAACCTCAAGCTGTGACCTGGAGAGCCAAATGAACCTGTCCTGCTGTCTCAGAGGGT 1092
QY 182 GLyGLnLYsSerGLuThrProLeuSerProProCysAlaAspSerAlaThrSerGluAla 201
DB 1093 GGACAGAAAAGCAGAGCGCCCTCACTCAGCCCTGTGCCGACTCTGCTACCTCAGAGCA 1152
QY 202 AlaGLyGLnArgProValHisProAlaSerValProSerValProSerValGLnProSer 221
DB 1153 GCAGGACAAAGCCCGTGTCATCCCGCAGCGGTGCCAGCGTCCAGCGGTGCGAGCGCTG 1212
QY 222 LeuLeuGLuAspSerProLeuValGLnAlaLeuSerGLyLeuHisLeuSerAlaAspArg 241
DB 1213 CTGTTAGAGCAGACGCCGCTGTGTACAGGGCTCAGTGGGCTGCACCTGTCCGACACAGG 1272
QY 242 LeuGLuAspSerAsnLYsIleLYsArgSerPheSerLeuAspIleLYsSerValSerTYr 261
DB 1273 CTGGAAGACAGCAATAAGCTCAAGCGTTCTCTCTCGATATCAATCAGTTTTCATAT 1332
QY 262 SerAlaSerMetAlaAlaSerLeuHisGLyPheSerSerSerGluAspAlaLeuGLuTYr 281
DB 1333 TCAGCCAGCATGGCAGCATCTTACATGGCTTCTCTCATCAGAGATGCTTTGGAATAC 1392
QY 282 TYrLYsProSerThrThrLeuAspGLyThrAsnLYsLeuCysGLnPheSerProValGLn 301
DB 1393 TACAAACCTTCCACTACTCTGGATGGGACCAACAGCTATGCGAGTTCTCCCTGTTTCA 1452
QY 302 GLuLeuSerGLuGLnThrProGLuThrSerProAspLYsGLuGLuAlaSerIleProLYs 321
DB 1453 GAACATTCGGAGCAGATCTCCGAAACAGTCTCTGATAGAGGAGGAGCCAGCATCCCCAAG 1512
QY 322 LYsLeuGLnThrAlaArgProSerAspSerGLnSerLYsArgLeuHisSerValArgThr 341
DB 1513 AAGTGTGAGACCGCAGCGCTTCAGACAGCAGCAGCAGATGTCATTCGGTCGAGAACC 1572
QY 342 SerSerSerGLyThrAlaGLnArgSerLeuLeuSerProLeuHisArgSerGLySerVal 361
DB 1573 AGCAGCATGTGCACCGCCAGAGGTCCCTTTATCTCCACTGATCGAAGTGGAGCGTG 1632
QY 362 GLuAspAsnTYrHisThrSerPheLeuPheGLyLeuSerThrSerGLnGLnHisLeuThr 381
DB 1633 GAGGACAAATTACCACACAGCGTTCCTTTTCGGCGTTTCCACCAGCCAGCAGCCTCAG 1692
```

```
QY 382 LYsSerAlaGLyLeuGLyLeuLYsGLyTYrPheHisSerAspIleLeuAlaProGLnThrSer 401
DB 1693 AAGTCTGTGGCTGGGCTTAAGGGCTGGCACTCGCATATCTTGGCCCCCAGACCTCT 1752
QY 402 ThrProSerLeuThrSerSerTYrPheAlaThrGluSerSerHisPheTYrSerAla 421
DB 1753 ACCCTTCCCTGACGAGCGTGTATTTTCCACAGAGTCTTCACTTCTACTCTGCTGCC 1812
QY 422 SerAlaIleTYrGLySerAlaSerTYrSerAlaTYrSerCysSerGLnLeuProThr 441
DB 1813 TCAGCCATCTACGGAGGCGAGTGCCAGTTACTCTGCTACAGCTGACGAGCTGCCACT 1872
QY 442 CysGLyAspGLnValTYrSerValArgArgGlnLYsProSerAspArgAlaAspSer 461
DB 1873 TCCGAGACCAAGTCTATTCTGTGCGCAGGCGGAGAGCCCAAGTGACAGCTGATCG 1932
QY 462 ArgArgSerTYrPheHisGLuGLuSerProPheGLuLYsGLnPheLYsArgArgSerCysGLn 481
DB 1933 CGCGGAGCTGGCATGAAGAGAGCCCTTTTGAAAGCAGTTTAAACGACAGAGCTGCCAA 1992
QY 482 MetGluPheGLyGluSerIleMetSerGluAsnArgSerArgGluLeuGLyLYsVal 501
DB 1993 ATGGAATTTGGAGAGAGCATCATGTCAAGAACAGGTCACGGGAAGAGCTGGGAAAGTG 2052
QY 502 GlySerGLnSerSerPheSerGlySerMetGluIleLeuGluValSer 517
DB 2053 GGCAGTCACTAGCTTTTTCGGCAGCATGGAATCATTTAGGCTCTCC 2100

RESULT 10
US-10-072-012-257
; Sequence 257, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zernusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Shinkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Esha
; APPLICANT: Padigaru, Muradidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grosse, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
```

;; PRIOR APPLICATION NUMBER: 60/266,767  
;; PRIOR FILING DATE: 2001-02-05  
;; PRIOR APPLICATION NUMBER: 60/267,057  
;; PRIOR FILING DATE: 2001-02-07  
;; PRIOR APPLICATION NUMBER: 60/266,975  
;; PRIOR FILING DATE: 2001-02-07  
;; PRIOR APPLICATION NUMBER: 60/267,459  
;; PRIOR FILING DATE: 2001-02-08  
;; Remaining Prior Application data removed - See File Wrapper or PALM.  
;; NUMBER OF SEQ ID NOS: 1391  
;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO 257  
;; LENGTH: 2071  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-10-072-012-257

## Alignment Scores:

Pred. No.: 7,166-297 Length: 2071  
Score: 304.00 Matches: 483  
Percent Similarity: 98.98% Conservative: 0  
Best Local Similarity: 98.98% Mismatches: 5  
Query Match: 98.80% Indels: 5  
DB: 12 Gaps: 0

US-09-964-277-21 (1-517) x US-10-072-012-257 (1-2071)

```
QY 30 GluLeuMetGlnGlnAsnGlyLeuGlyTyrValLeuAsnAlaSerThrCysProLys 49
DB 592 GAGCTGATGCGAGCAATGGGATTTGGTTATGTGTTAAATCCAGCAATACCTGTCCAAAG 651
QY 50 ProAspPheIleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGlu 69
DB 652 CCGTACTTATCCCGAGTCTCAATTCCTCGCTGCTGCTGTAATCAGACTTTTGTGAG 711
QY 70 LysIleLeuProTrpLeuAspLysSerValAspPheIleGluLysAlaSerAsn 89
DB 712 AAAATTTTGGCGGTGTTGGCAAAATCAGTAGATTTCATTGAGAAAGCAAAAGCCTCCAAAT 771
QY 90 GlyCysValLeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAla 109
DB 772 GGATGTTTCTAGTGCACTGTTAGCTGGGATCTCCCGCTCCGCCACCATCGCTATCGCC 831
QY 110 TyrIleMetLysArgMetAspMetSerLeuAspGluAlaTyrArgPheValLysGluLys 129
DB 832 TACATCATGAAGAGATGGACATGCTTTTACATGAAGCTTACAGATTGTGAAAGAAAA 891
QY 130 ArgProThrIleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGluLysLys 149
DB 892 AGACCTACTATATCTCCAAACTTCNAATTTCTGGGCCAACTCTCTGGACTATGAGAGAG 951
QY 150 IleLysAsnGlnThrGlyAlaSerGlyProLysSerLysLysLeuLysLeuLeuHisLeuGlu 169
DB 952 ATTAAGAACCAAGCTGGAGCATCAGGGCAAAAGCAAACTCAAGCTGCTGCACCTGGAG 1011
QY 170 LysProAsnGluProValProAlaValSerGluGlyGlyGlnLysSerGluThrProLeu 189
DB 1012 AAGCCAAATGAACCTGCTCCCTGCTGCTCAGAGGTGACAGAAAAGCAGACGCCCTC 1071
QY 190 SerProProCysAlaAspSerAlaThrSerGluAlaAlaGlyGlnArgProValHisPro 209
DB 1072 AGTCACCTCTGGCCGACTGCTACCTCAGAGCAGCAGCAAAAGCCCGCTGCATC-- 1129
QY 210 AlaSerValProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuVal 229
DB 1130 C--C--G--CC--AGCGTACCAGCGTGCAGCGCTGCTGTTAGAGGACAGCCCGCTGGTA 1182
QY 230 GlnAlaLeuSerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnLysLeuLys 249
DB 1183 CAGGGGCTCAGTGGGCTGCACCTGTCCGACACAGGCTGGAGACAGCAATAAGCTCAAG 1242
QY 250 ArgSerPheSerLeuAspIleLysSerValSerTyrSerAlaSerMetAlaAlaSerLeu 269
```

```
DB 1243 CGTTCCTTCTCTCGGATATCAAAATCAGTTTTCATATTACCCAGCATGGCAGCATCCTTA 1302
QY 270 HisGlyPheSerSerSerGluAspAlaLeuGluTyrTyrLysProSerThrThrLeuAsp 289
DB 1303 CATGGCTTCTCC*CATCAGAAAGATGCTTTGGAATACTACAAACCTTCCACTACTCTGGAT 1362
QY 290 GlyThrAsnLysLeuCysGlnPheSerProValGlnGluLeuSerGluGlnThrProGlu 309
DB 1363 GGGACCAACAAAGCTATGCCAGTTCTCCCTGTTCAGAACTATCGAGAGCAGACTCCCGAA 1422
QY 310 ThrSerProAspLysGluGluAlaSerIleProLysLysLeuGlnThrAlaArgProSer 329
DB 1423 ACCAGTCTCTGATAAGGAGGAAGCCAGCATCCCAAGAAGCTGCAGACCCAGCGCTTCA 1482
QY 330 AspSerGlnSerLysArgLeuHisSerValArgThrSerSerGlyThrAlaGlnArg 349
DB 1483 GACAGCCAGAGCAGAGCATTTGCTTCAGAACCCAGCAGCAGTGGCAGCCGCCAGAGG 1542
QY 350 SerLeuLeuSerProLeuHisArgSerGlySerValGluAspAsnTyrHisThrSerPhe 369
DB 1543 TCCCTTTTATCTCCACTGTCATCGAAGTGGGAGCGTGGAGCAAT*ACCAACACAGCTTC 1602
QY 370 LeuPheGlyLeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuLys 389
DB 1603 CTTTTCGGCCTTTCCACCAGCCAGCACCTCACGAAGTCTGCTGGCCTGGGCCCTTAAG 1662
QY 390 GlyTrpHisSerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTrp 409
DB 1663 GGCTGGCAGCTCGGATATCTTGGCCCCCAGACCTCTACCCCTTCCCTGACCAGCAGCTGG 1722
QY 410 TyrPheAlaThrGluSerSerHisPheTyrSerAlaSerAlaIleTyrGlyGlySerAla 429
DB 1723 TATTTGCCACAGAGTCTCACACTTCTACTCTGCTCAGCCATCTACGAGGAGCAGTGCC 1782
QY 430 SerTyrSerAlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerVal 449
DB 1783 AGTTACTCTGCTTACAGCTGCAGCCAGCTGCCCACTTGGGAGACCAAGTCTATTCTGTG 1842
QY 450 ArgArgArgGlnLysProSerAspArgAlaAspSerArgArgSerTrpHisGluGluSer 469
DB 1843 CGCAGCGCGGAGAGCCCAAGTGCAGAGCTGACTCGCGCGGAGCTGGCATGAAGAGAGC 1902
QY 470 ProPheGluLysGlnPheLysArgArgSerCysGlnMetGluPheGlyGluSerIleMet 489
DB 1903 CCCTTGAAAGAGAGTTTAAACGCAAGAGCTGCCAATGGAAATTTGGAGAGCATCATG 1962
QY 490 SerGluAsnArgSerArgGluGluLeuGlyLysValGlySerGlnSerSerPheSerGly 509
DB 1963 TCAGAGAACAGGTCACGGGAAGAGCTGGGGAAGTGGGCAAGTGGCAGTCAGTCTAGCTTTTCGGGC 2022
QY 510 SerMetGluIleIleGluValSer 517
DB 2023 AGCATGGAAATCATTTGAGGTCTCC 2046
```

## RESULT 11

US-10-108-260A-2429  
; Sequence 2429, Application US/10108260A  
; Publication No. US20040005560A1  
; GENERAL INFORMATION:  
; APPLICANT: HELIX RESEARCH INSTITUTE  
; TITLE OF INVENTION: NO. US20040005560A1e1 full length cDNA  
; FILE REFERENCE: H1-A0106  
; CURRENT APPLICATION NUMBER: US/10/108,260A  
; CURRENT FILING DATE: 2002-03-27  
; NUMBER OF SEQ ID NOS: 5458  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2429  
; LENGTH: 1916  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-108-260A-2429

Alignment Scores:

```

Pred. No.: 1,32e-169 Length: 1916
Score: 178.00 Matches: 278
Percent Similarity: 99.29% Conservative: 0
Best Local Similarity: 99.29% Mismatches: 1
Query Match: 34.43% Indels: 2
DB: 15 Gaps: 0

US-09-964-277-21 (1-517) x US-10-108-260A-2429 (1-1916)

Qy 239 AlaAspArgLeuGluAspSerLeuLysArgSerPheSerLeuAspIleLysSer 258
Db 142 GCAGACAGGCTGGAAGACAGCAATAGCTCAAGCGTCTCTCTCTCGATATCAATCA 201
Qy 259 ValSerTyrSerAlaSerMetAlaSerLeuHisGlyPheSerSerGluAspAla 278
Db 202 GTTTCATATTCACCCAGCATGGCAGCATCTTACATGGCTTCTCTCATCAGATGCT 261
Qy 279 LeuGluTyrTrpLysProSerThrThrLeuAspGlyThrAsnLysLeuCyGlnPheSer 298
Db 262 TTGGAATACCTACAAAGCTTCCACTCTCTGGATGGACCAACAGCTATGCCAGTTCTCC 321
Qy 299 ProValGlnGluSerGluGlnThrProGluThrSerProAspLysGluGluAlaSer 318
Db 322 CCGTTCAGGAATATCGGAGAGATCCCGAACCAGTCTCTGATAGAGAGAGCCAGC 381
Qy 319 IleProLysLysLeuGlnThrAlaArgProSerAspSerGlnSerLysArgLeuHisSer 338
Db 382 ATCCCAAGAAGCTGCAGACCGCCGCTTCAGACAGCCAGCAAGCGATTGCATTGC 441
Qy 339 ValArgThrSerSerGlyThrAlaGlnArgSerLeuLeuSerProLeuHisArgSer 358
Db 442 GTGCAACACAGCAGAGTGCCACCGCCGAGAGTCCCTTTATCTCCACTGATCGAAGT 501
Qy 359 GlySerValGluAspAsnTyrHisThrSerPheLeuPheGlyLeuSerThr-SerGlnG 378
Db 502 GGGAGCGTGGAGACATATACACACAGCTCTTTCGGGCTTTCAT-CAGCCAGCA 560
Qy 378 nHisLeuThrLysSerAlaGlyLeuGlyLeuLysGlyTrpHisSerAspIleLeuAlaPr 398
Db 561 GCACCTCACGAAGTCTGCTGGCGTGGGCTTAAGGGCTGGCACTCGGATATCTTGGCCCC 620
Qy 398 oGlnThrSerThr-ProSerLeuThrSerSerTrpTyrPheAlaThrGluSerSerHisph 418
Db 621 CCAGACCTTACCCCTTCCCTGACAGCAGCTGTATTTTGGCCACAGAGTCTCACACT 680
Qy 418 eTyrSerAlaSerAlaIleTyrGlyGlySerAlaSerTyrSerAlaTyrSerCysSerGl 438
Db 681 CTACTCTGCTCAGCCATCTACGGAGGAGTGCCTTACTCTGCTACAGCTGCAGCCA 740
Qy 438 nLeuProThrCysGlyAspGlnValTyrSerValArgArgGlnLysProSerAspAr 458
Db 741 GCTGCCACCTTGGGAGACCAAGTCTATCTGTGCGAGCGGAGGAGGAGGAGGAG 800
Qy 458 gAlaAspSerArgArgSerTrpHisGluGluSerProPheGluLysGlnPheLysArgAr 478
Db 801 AGCTGACTCGCGGCGAGCTGGCATGAAGAGAGCCCTTTGAAAGAGCAGTTTAAACGCG 860
Qy 478 gSerCysGlnMetGluPheGlyGluSerIleMetSerGluAsnArgSerArgGluGluLe 498
Db 861 AAGCTGCAATGGAAATTTGGAGAGAGCATCATGTGAGAGACAGGTACCGGGAAGGCT 920
Qy 498 uGlyLysValGlySerGlnSerPheSerGlySerMetGluIleIleGluValSer 517
Db 921 GGGGAAGTGGCAGTCACTGCTAGCTTTTCGGGAGCATGGAAATCATTTAGGTCTCC 978

```

RESULT 12

US-09-783-590-8120  
 ; Sequence 8120, Application US/09783590  
 ; Patent No. US20020110850A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Dillon, Patrick J.  
 ; APPLICANT: Haseltine, William A.  
 ; APPLICANT: Li, Haodong

```

; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
; FILE REFERENCE: PO-16.2C1
; CURRENT APPLICATION NUMBER: US/09/783,590
; CURRENT FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 08/420,856
; PRIOR FILING DATE: 1995-04-12
; PRIOR APPLICATION NUMBER: 08/346,731
; PRIOR FILING DATE: 1994-11-21
; NUMBER OF SEQ ID NOS: 12485
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8120
; LENGTH: 424
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (10)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (65)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (128)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (167)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (199)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (237)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (246)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (250)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (255)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (272)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (277)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (290)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (291)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (292)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (298)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (300)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (317)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (328)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (331)

```



OTHER INFORMATION: n equals a,t,g, or c  
NAME/KEY: misc feature  
LOCATION: (335)  
OTHER INFORMATION: n equals a,t,g, or c  
NAME/KEY: misc feature  
LOCATION: (365)  
OTHER INFORMATION: n equals a,t,g, or c  
NAME/KEY: misc feature  
LOCATION: (395)  
OTHER INFORMATION: n equals a,t,g, or c  
NAME/KEY: misc feature  
LOCATION: (403)  
OTHER INFORMATION: n equals a,t,g, or c  
NAME/KEY: misc feature  
LOCATION: (404)  
OTHER INFORMATION: n equals a,t,g, or c  
NAME/KEY: misc feature  
LOCATION: (413)  
OTHER INFORMATION: n equals a,t,g, or c  
US-09-783-590-8120

Alignment Scores:  
Pred. No.: 6.7e-43 Length: 424  
Score: 52.00 Matches: 52  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 10.06% Indels: 0  
DB: 9 Gaps: 0

US-09-964-277-21 (1-517) x US-09-783-590-8120 (1-424)

QY 389 LysGlyTrpHisSerAspLeuAlaProGlnThrSerThrProSerLeuThrSerSer 408  
DB 15 AAGGCTGGCATCGGATATCTTGGCCCCCAGACCTTACCCCTTCCCTNACGAGC 74  
QY 409 TrpTyrPheAlaThrGlnSerSerHisPheTyrSerAlaSerAlaLeuTyrGlySer 428  
DB 75 TGGTATTTTGGCAGAGAGTCTTACACTTCTACTGCTCAGCCATCTAGGNGGAGT 134  
QY 429 AlaSerTyrSerAlaTyrSerCysSerGlnLeuPro 440  
DB 135 GCAGTTACTTGGCTTACAGCTGCAGCCAGCTNCCC 170

RESULT 13  
US-10-005-858-1  
Sequence 1, Application US/10005858  
Publication No. US20020116729A1  
GENERAL INFORMATION:  
APPLICANT: Allen, Keith D.  
TITLE OF INVENTION: TRANSGENIC MICE CONTAINING NTTP1  
FILE REFERENCE: R-690  
CURRENT APPLICATION NUMBER: US/10/005,858  
CURRENT FILING DATE: 2001-12-04  
PRIOR APPLICATION NUMBER: US 60/251,802  
PRIOR FILING DATE: 2000-12-06  
NUMBER OF SEQ ID NOS: 3  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1  
LENGTH: 2453  
TYPE: DNA  
ORGANISM: Mus musculus  
US-10-005-858-1

Alignment Scores:  
Pred. No.: 6.96e-10 Length: 2453  
Score: 20.00 Matches: 20  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 3.87% Indels: 0  
DB: 13 Gaps: 0

US-09-964-277-21 (1-517) x US-10-005-858-1 (1-2453)

QY 94 ValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMetLys 113  
DB 827 GTTCACTGTCTGGCTGGCATCTCTCGTCTGCCACCATTCGCCATCATCATGAAA 886

## RESULT 14

US-09-964-824A-463/c  
Sequence 463, Application US/09964824A  
Patent No. US20020102531A1  
GENERAL INFORMATION:  
APPLICANT: Horrigan, Stephen  
TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu  
FILE REFERENCE: 689290-73  
CURRENT APPLICATION NUMBER: US/09/964,824A  
CURRENT FILING DATE: 2001-09-27  
PRIOR APPLICATION NUMBER: US/60/236,033  
PRIOR FILING DATE: 2000-09-28  
PRIOR APPLICATION NUMBER: US/60/236,032  
PRIOR FILING DATE: 2000-09-28  
PRIOR APPLICATION NUMBER: US/60/236,028  
PRIOR FILING DATE: 2000-09-28  
NUMBER OF SEQ ID NOS: 583  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 463  
LENGTH: 263  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1)...(263)  
OTHER INFORMATION: n=a,t,g or c  
US-09-964-824A-463

Alignment Scores:  
Pred. No.: 0.0112 Length: 263  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 2.32% Indels: 0  
DB: 9 Gaps: 0

US-09-964-277-21 (1-517) x US-09-964-824A-463 (1-263)

QY 92 ValLeuValHisCysLeuAlaGlyIleSerArgSer 103  
DB 105 GTCCTGTACATTCCTGCTGGCATTAGCGNTCA 70

## RESULT 15

US-09-880-107-1814/c  
Sequence 1814, Application US/09880107  
Patent No. US20020142981A1  
GENERAL INFORMATION:  
APPLICANT: Horne, Darci T.  
APPLICANT: Vockley, Joseph G.  
APPLICANT: Scherf, Uwe  
APPLICANT: Gene Logic, Inc.  
TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer  
FILE REFERENCE: 44921-5028-WO  
CURRENT APPLICATION NUMBER: US/09/880,107  
CURRENT FILING DATE: 2001-06-14  
PRIOR APPLICATION NUMBER: US 60/211,379  
PRIOR FILING DATE: 2000-06-14  
PRIOR APPLICATION NUMBER: US 60/237,054  
PRIOR FILING DATE: 2000-10-02  
NUMBER OF SEQ ID NOS: 3950  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1814  
LENGTH: 263  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: Genbank Accession No. US20020142981A1 F10276

```
; NAME/KEY: unsure
; LOCATION: (1)..(263)
; OTHER INFORMATION: n = a or c or g or t
US-09-880-107-1814

Alignment Scores:
Pred. No.:      0.0112      Length:      263
Score:          12.00       Matches:      12
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:      2.32%      Indels:      0
DB:               9         Gaps:        0

US-09-964-277-21 (1-517) x US-09-880-107-1814 (1-263)

Qy      92 ValIeuValHisCysIeuAlaGlyIleSerArgSer 103
Db      105 GTCTTGTCATTGCTTGGCTGGCATTAGCCGNICA 70

Search completed: February 28, 2004, 04:23:22
Job time : 532 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 28, 2004, 00:54:04 ; Search time 3854 Seconds  
(without alignments)  
4005.902 Million cell updates/sec

Title: US-09-964-277-21  
Perfect score: 517  
Sequence: 1 MFLSLQTVSLFYFWNRR.....LGKVGSSFSFGSMIEIVS 517

Scoring table: OLIGO  
Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Word size: 1

Total number of hits satisfying chosen parameters: 55025477

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters: -DEV-xml  
-MODEU-frame\_p2n.model -DEV-xml  
-Q/cgn2\_1/USPTO\_spool/SUSF0984277/runat\_25022004\_085401\_2218/app\_query.fasta\_1.711  
-DB=EST -QFMT=fastap -SUFFIX=olip2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto  
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US0984277@cgn.1.1.2810 -runat\_25022004\_085401\_2218 -NCPU=6 -ICPU=3  
-NO\_WMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -Fgapop=6  
-Fgapext=7 -Ygapop=60 -Ygapext=6 -DELOP=6 -DELEXT=7

Database :

- 1: em estba.\*
- 2: em esthum.\*
- 3: em estm.\*
- 4: em estmu.\*
- 5: em estov.\*
- 6: em estpl.\*
- 7: em estro.\*
- 8: em htc.\*
- 9: gb est1.\*
- 10: gb est2.\*
- 11: gb htc.\*
- 12: gb est3.\*
- 13: gb est4.\*
- 14: gb est5.\*
- 15: em estfun.\*
- 16: em estom.\*
- 17: em gss hum.\*
- 18: em gss inv.\*
- 19: em gss pln.\*
- 20: em gss vrt.\*
- 21: em gss fun.\*
- 22: em gss nam.\*
- 23: em gss mus.\*
- 24: em gss pro.\*
- 25: em gss rod.\*
- 26: em gss phg.\*
- 27: em gss vrl.\*
- 28: gb gss1.\*

29: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	216	41.8	903	12	BG482429	BG482429 602526595
2	209	40.4	891	13	BQ670989	BQ670989 AGENCOURT
3	186	36.0	1060	12	BM546940	BM546940 AGENCOURT
4	179	34.6	983	12	BG165188	BG165188 602344051
5	167	32.3	859	13	BQ933499	BQ933499 AGENCOURT
6	160	30.9	881	13	BQ930140	BQ930140 AGENCOURT
7	160	30.9	974	13	BQ951695	BQ951695 AGENCOURT
8	159	30.8	663	13	BX479029	BX479029 DKFZP686J
9	139	26.9	883	13	BQ945892	BQ945892 AGENCOURT
10	130	25.1	391	13	EX475224	EX475224 DKFZP686J
11	126	24.4	444	12	BM987142	BM987142 UI-H-COO-
12	126	24.4	920	10	BE897795	BE897795 601438457
13	124	24.0	621	9	AV703072	AV703072 AV703072
14	124	24.0	693	9	AV701628	AV701628 AV701628
15	101	19.5	660	12	BI917706	BI917706 603183686
16	96	18.6	463	10	BG058779	BG058779 raf10f04
17	96	18.6	682	12	BI816954	BI816954 imageqc_1
18	87	16.8	752	13	BQ432082	BQ432082 AGENCOURT
19	86	16.6	424	10	AW461438	AW461438 BP230005A
20	77	14.9	913	13	BU170187	BU170187 AGENCOURT
21	73	14.1	371	13	BX492954	BX492954 DKFZP781F
22	64	12.4	337	14	CB606296	CB606296 AMGNNUC-S
23	64	12.4	442	14	CB606149	CB606149 AMGNNUC-N
24	63	12.2	402	29	CG499793	CG499793 OST42180
25	63	12.2	404	10	BB16652	BB16652 BB16652
26	63	12.2	412	9	AA879894	AA879894 VW01D07.r
27	63	12.2	421	10	BE946894	BE946894 UI-M-BH3-
28	63	12.2	579	10	BF582366	BF582366 602094304
29	63	12.2	715	14	CF727177	CF727177 UI-M-HB0-
30	63	12.2	769	13	BU704078	BU704078 UI-M-FOO-
31	63	12.2	836	12	BI821804	BI821804 603035893
32	63	12.2	862	13	BQ770036	BQ770036 UI-M-F10-
33	63	12.2	946	12	BI657528	BI657528 603281866
34	61	11.8	386	10	BE983039	BE983039 UI-M-CGQ0
35	61	11.8	476	28	AZ561697	AZ561697 RPCI-23-1
36	61	11.8	580	13	BQ568918	BQ568918 g118d07
37	61	11.8	634	14	CF733313	CF733313 UI-M-HB0-
38	61	11.8	680	13	BU613778	BU613778 UI-M-EMO-
39	61	11.8	723	13	BU708175	BU708175 UI-M-FCO-
40	61	11.8	1110	10	BF143207	BF143207 601787214
41	54	10.4	736	29	CE172184	CE172184 tigr-gss-
42	53	10.3	463	9	AA791519	AA791519 vs39g09.r
43	51	9.9	655	14	CF532917	CF532917 UI-M-GHO-
44	45	8.7	743	14	CF727802	CF727802 UI-M-HB0-
45	44	8.5	639	14	CF742387	CF742387 UI-M-HB0-

ALIGNMENTS

RESULT 1  
BG482429  
LOCUS  
DEFINITION  
602526595F1 NIH\_MGC\_21 Homo sapiens cDNA clone IMAGE:4650430 5',  
mRNA sequence.  
ACCESSION  
BG482429  
VERSION  
BG482429.1 GI:13414708  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 903)

441 AAGCAAAATGAACCTGCTCCTCTGCTCTCAGAGGGTGGACAGAAAAAGCAGACGCCCTC 500

190 SerProProCysAlaAaspSerAlaThrSerGluAlaAlaGlyGlnAArgProValHisPro 209

501 AGTCCACCTCTGTGCCCACTCTGCTACCTCAGAGGGCAGCAGGACAAAGCCCGCTGCATCCC 560

210 AlaSerValProSerValProSerValGlnProSerLeuLeuGluAAspSerProLeuVal 229

561 GCCAGCGTGGCCAGCGTGCCACGCTGCAGCGCTGCCTGCTT\*AGAGGACAGCCCGCTGGTA 620

230 GlnAlaLeuSerGlyLeuHisLeuSerAlaAaspArgLeuGluAAspSer 245

621 CAGCGCGTCAGTGGGCTGCACCTGTCCGACAGCAGCGCTGGAAGACAGC 668

RESULT 2

BQ670989

LOCUS

DEFINITION BQ670989 891 bp mRNA linear EST 15-JUL-2002

5', mRNA sequence. NIH\_MGC\_102 Homo sapiens cDNA clone IMAGE:62115899

ACCESSION BQ670989

VERSION BQ670989.1 GI:21781823

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS 1. (bases 1 to 891)

TITLE NIH-MGC <http://mgc.nci.nih.gov/>.

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: [cgabbs-x@mail.nih.gov](mailto:cgabbs-x@mail.nih.gov)

Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Plate: LICM2374 row: f column: 06

High quality sequence stop: 633.

FEATURES

source

1..891

Location/Qualifiers

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clones="IMAGE:6211589"

/tissue\_type="epidermoid carcinoma, cell line"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="NIH\_MGC\_102"

/note="Organ: salivary gland; Vector: pOTB7; Site\_1: XhoI; Site\_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.:	6.48e-184	Length:	891
Score:	209.00	Matches:	209
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	40.43%	Indels:	0
DB:	13	Gaps:	0

US-09-964-277-21 (1-517) x BQ670989 (1-891)

Qy 309 GluThrSerProAspLysGluGluAlaSerIleProLysLeuGlnThrAlaArgPro 328

Db 1 GAAACCACTCTGATAGGAGGAGCCAGCATCCCAAGAGCTCGACACGCCAGCGCT 60

```
QY 329 SerAspSerGlnSerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGln 348
Db 61 TCAGACAGCCAGAGCAAGCATTCGATTCGGTCAGAACCCAGCAGCAGTGGCAGCCGAC 120
QY 349 ArgSerLeuLeuSerProLeuHisArgSerGlySerValGluAspAenTyrHisThrSer 368
Db 121 AGGTCCCTTTTATCTCCACTGCATCGAAGTGGAGCGTGGAGGACAATACCCACACAGC 180
QY 369 PheLeuPheGlyLeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLeu 388
Db 181 TTCCTTTTCGGCCTTTCCACAGCCAGCAGCAGCAGCTCAGGAAGTCTGCTGGCCTGGGCTT 240
QY 389 LysGlyTyrHisSerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSer 408
Db 241 AAGGGCTGGCACTCGATATCTTGCCCCCAGACCTTACCCTTCCCTGGACACAGCAGC 300
QY 409 TrpTyrPheAlaThrGluSerSerHisPheTyrSerAlaSerAlaIleTyrGlyGlySer 428
Db 301 TGGTATTTTGGCACAGAGTCTCTACACTTCTACTCTGCTCAGCCATCTACGGAGCAGT 360
QY 429 AlaSerTyrSerAlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSer 448
Db 361 GCCAGTTACTCTGCTCAGAGTGGAGCCAGCTGCCACTTGGGAGACCAAGCTATTCT 420
QY 449 ValArgArgArgGlnLysProSerAspArgAlaAspSerArgArgSerTrpHisGlyGlu 468
Db 421 GTGCCAGGGCGCAGAACCCAGGTGACAGAGCTGCTCGCGCGAGCTGGCATGAGAG 480
QY 469 SerProPheGlyLysGlnPheLysArgArgSerCysGlnMetGluPheGlyGluSerIle 488
Db 481 AGCCCCCTTTGAAAGCAGTCTTAAACGCCAGAGAGTCCCAATGGAATTTGGAGAGAGCATC 540
QY 489 MetSerGluAsnArgSerArgGluGluLeuGlyLysValGlySerGlnSerSerPheSer 508
Db 541 ATGTGAGAGACAGGTACCGGAGAGAGTGGGAAAGTGGGCAGTCACTAGCTAGCTTTTCG 600
QY 509 GlySerMetGluIleIleGluValSer 517
Db 601 GCACGATGGAATCATTTAGGCTCTCC 627
```

```
RESULT 3
BM546940
LOCUS
DEFINITION 6491295 NIH_MGC_125 Homo sapiens cDNA clone IMAGE:5723711
5', mRNA sequence.
```

```
ACCESSION BM546940
VERSION BM546940.1 GI:18780310
```

```
KEYWORDS EST.
```

```
SOURCE Homo sapiens (human)
```

```
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

```
REFERENCE 1 (bases 1 to 1060)
```

```
AUTHORS NIH-MGC http://mrc.nci.nih.gov/.
```

```
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
```

```
JOURNAL Unpublished (1999)
```

```
COMMENT Contact: Robert Strausberg, Ph.D.
```

```
Email: cgabbs-remail.nih.gov
```

```
Tissue Procurement: Invitrogen
```

```
cDNA Library Preparation: Life Technologies, Inc.
```

```
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
```

```
Clone distribution by: Agencourt Bioscience Corporation
```

```
found through the I.M.A.G.E. Consortium/LLNL at:
```

```
http://image.llnl.gov
```

```
Plate: LLAM12711 row: m column: 24
```

```
High quality sequence stop: 737.
```

```
1. .1060
```

```
location/Qualifiers
```

```
FEATURES
```

```
source
```

```
/organism="Homo sapiens"
```

```
/mol_type="mRNA"
```

```
/db_xref="taxon:9606"
```

```
RESULT 4
```

```
BM546940
```

```
LOCUS
```

```
DEFINITION 602344051F1 NIH_MGC_89 Homo sapiens cDNA clone IMAGE:4454029 5',
```

```
mRNA sequence.
```

```
ACCESSION BM546940
```

```
VERSION BM546940.1 GI:12671891
```

```
KEYWORDS EST.
```

```
SOURCE Homo sapiens (human)
```

```
/clone="IMAGE:5723711"
```

```
/lab_host="DH10B"
```

```
/clone_lib="NIH_MGC_125"
```

```
/note="Organ: ovary (pool of 3); Vector: pCMV-SPORT6;
```

```
Site 1: EcoRV (destroyed); Site 2: NotI; RNA source pool
```

```
of three ovaries, from females ranging in age from 38 to
```

```
49 yo. Library is oligo-dT primed and directionally cloned
```

```
(EcoRV site is destroyed upon cloning). Average insert
```

```
size 2.1 kb, insert size range 1-3.5 kb. Library is
```

```
normalized and enriched for full-length clones and was
```

```
constructed by C. Gruber (Invitrogen). Research Genetics
```

```
tracking code 036."
```

```
ORIGIN
```

```
Alignment Scores:
```

```
Pred. No.: 2,13e-162 Length: 1060
```

```
Score: 186.00 Matches: 186
```

```
Percent Similarity: 100.00% Conservative: 0
```

```
Best Local Similarity: 100.00% Mismatches: 0
```

```
Query Match: 35.98% Indels: 0
```

```
DB: 12 Gaps: 0
```

```
US-09-964-277-21 (1-517) x BM546940 (1-1060)
```

```
QY 30 GluLeuMetGlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerAsnThrCysProLys 49
```

```
Db 51 CAGCTGATGCAGCAGATGGGATTGGTTATGTGTTAAATGCCAGCAATACCTGTCCAAAG 110
```

```
QY 50 ProAspPheIleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGlu 69
```

```
Db 111 CCGTACITTTATCCCGAGTCTCATTTCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 170
```

```
QY 70 LysIleLeuProTyrLeuAspLysSerValAspPheIleGlyLysAlaLysAlaSerAsn 89
```

```
Db 171 AAAATTTTGGCTGGTGGCAATCATAGATTTCTATTGAGAAAGCAAAAGCTCCCAT 230
```

```
QY 90 GlyCysValLeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAla 109
```

```
Db 231 GGATGTGTTCTAGTGCACCTGTTTAGCTGGGATCTCCCGCTCCGCCACCATCGCTATCGCC 290
```

```
QY 110 TyrIleMetLysArgMetAspMetSerLeuAspGluAlaTyrArgPheValLysGluLys 129
```

```
Db 291 TACATCATGAGAGATGGACATGTCCTTTAGTAGTAGCTTACAGATTTGTGAAAGGAAAA 350
```

```
QY 130 ArgProThrIleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGluLysLys 149
```

```
Db 351 AGACCTACTATATCTCAAACTTCAATTTTCTGGCCCACTCTCTGGACTATGAGAAAGAG 410
```

```
QY 150 IleLysAsnGlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuLeuHisLeuGlu 169
```

```
Db 411 ATTAAGAACCAAGCTGGAGCATCAGGCCCAAGAGCAAACTCAAGCTGTGACCTGGAG 470
```

```
QY 170 LysProAsnGluProValProAlaValSerGluGlyGlnLysSerGluThrProLeu 189
```

```
Db 471 AAGCAAAATGAACCTGCTCCCTGCTGCTCTCAGAGGGTGGACAGAAAGCAGACGCCCTC 530
```

```
QY 190 SerProProCysAlaAspSerAlaThrSerGluAlaAlaGlyGlnArgProValHisPro 209
```

```
Db 531 AGTCCACCTGTGCGGACTCTGTACCTACCTCAGAGCAGGACAAAGGCCGCCGTGCATCC 590
```

```
QY 210 AlaSerValProSerVal 215
```

```
Db 591 GCCAGCTACCCAGCGTG 608
```

```
BM546940 983 bp mRNA linear EST 06-FEB-2001
```

```
602344051F1 NIH_MGC_89 Homo sapiens cDNA clone IMAGE:4454029 5',
```

```
mRNA sequence.
```

```
ACCESSION BM546940
```

```
VERSION BM546940.1 GI:12671891
```

```
KEYWORDS EST.
```

```
SOURCE Homo sapiens (human)
```

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 983)  
 AUTHORS NIH-MGC http://hgnc.ncl.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-r@mail.nih.gov  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LHAM10245 row: f column: 14  
 High quality sequence stop: 625.  
 Location/Qualifiers  
 FEATURES  
 source  
 1. 983  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4454029"  
 /tissue\_type="hypernephroma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH MGC 89"  
 /notes="Organ: kidney; Vector: pCMV-SPORT6; Site 1: NotI;  
 Site 2: SalI; Cloned unidirectionally; oligo-dT primed.  
 Average insert size 1.3 kb. Library enriched for  
 full-length clones and constructed by Life Technologies.  
 Note: this is a NIH MGC Library."  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 6 6e-156 Length: 983  
 Score: 179.00 Matches: 192  
 Percent Similarity: 99.48% Conservative: 0  
 Best Local Similarity: 99.48% Mismatches: 0  
 Query Match: 34.62% Indels: 1  
 DB: 12 Gaps: 0  
 US-09-964-277-21 (1-517) x BG165188 (1-983)  
 Qy 271 GlyPheSerSerGluAlaLeuGluTyrTyrPheProSerThrThrLeuAspGly 290  
 Db 2 GGCCTCTCTCCATCAGAAGATGCTTTGGATACACAACTTCCACTACTCGATGG 61  
 Qy 291 ThrAsnLysLeuGlnPheSerProValGlnGluLeuSerGluGlnThrProGluThr 310  
 Db 62 ACCAACAGCTATGCCAGTTCTCCCTGTTTCAGAACTATCGAGCAGACTCCGAAACC 121  
 Qy 311 SerProAspLysGluGluAlaSerIleProLysLeuGlnThrAlaArgProSerAsp 330  
 Db 122 AGTCCTGATAGGAGGAAGCCAGCATCCCAAGCTGCAGACTCCAGGCTTCAGAC 181  
 Qy 331 SerGlnSerLysArgLeuHisSerValArgThrSerSerGlyThrAlaGlnArgSer 350  
 Db 182 AGCCAGAGCAAGCATTCGATTCGTCAGAACAGCAGCAGTGGCACCCTCCAGAGGTC 241  
 Qy 351 LeuLeuSerProLeuHisArgSerGlySerValGluAspAsnThrHisThrPheLeu 370  
 Db 242 CTTTTATCTCCATGTCATCGAAGTGGGAGCGTGGAGGACAAATTACACACAGCTCC 301  
 Qy 371 PheGlyLeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuLysGly 390  
 Db 302 TTCGGCTCTTCCACGACCGAGCAGCCTCAGCAAGTCTGCTGGCTGGCTTAAGGCG 361  
 Qy 391 TrpHisSerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTrpTyr 410  
 Db 362 TGGCACTCGGATATCTTGGGCCCCCAGACCTCTACCCCTTCCCTGACACGAGCTGT 421  
 Qy 411 PheAlaThrGluSerSerHisPheTyrSerAlaSerAlaIleTyrGlyGlySerAlaSer 430

Db 422 TTTCGCACAGAGTCTCTACACTTCTACTCTGCTCAGCATCTACGGAGGAGTCCAGT 481  
 Qy 431 TyrSerAlaTyrSerCysSerGlnLeuProThrCysGlyAspGln-ValTyrSerValAr 450  
 Db 482 TACTCTGCTACAGCTGCGAGCAGTCCCTTGGCGAGACCAAGTCTATTCTGTGG 541  
 Qy 450 GARGARGGlnLysProSerAspArgAlaAspSerArg 462  
 Db 542 CAGGCGGAGAGACCCAGTGCAGAGCTGACTCGGG 578  
 RESULT 5  
 LOCUS BQ933499  
 DEFINITION AGCOURT 8732489 NIH\_MGC\_101 Homo sapiens cDNA clone IMAGE:6455595  
 5' mRNA sequence.  
 ACCESSION BQ933499  
 VERSION BQ933499.1 GI:22348882  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 859)  
 AUTHORS NIH-MGC http://hgnc.ncl.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-r@mail.nih.gov  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLC2634 row: i column: 04  
 High quality sequence stop: 727.  
 Location/Qualifiers  
 FEATURES  
 source  
 1. 859  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6455595"  
 /tissue\_type="epidermoid carcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH MGC\_101"  
 /note="Organ: lung; Vector: pOTS7; Site 1: EcoRI; Site 2:  
 XhoI; cDNA made by oligo-dT priming. Directionally cloned  
 into EcoRI/XhoI sites using the following 5' adaptor:  
 GGCACGAG(G) Library constructed by Ling Hong in the  
 laboratory of Gerald M. Rubin (University of California,  
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
 Superscript II RT (Life Technologies). Note: this is a  
 NIH\_MGC Library."  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 8 8e-145 Length: 859  
 Score: 167.00 Matches: 267  
 Percent Similarity: 99.26% Conservative: 0  
 Best Local Similarity: 99.26% Mismatches: 1  
 Query Match: 32.30% Indels: 2  
 DB: 13 Gaps: 0  
 US-09-964-277-21 (1-517) x BQ933499 (1-859)  
 Qy 118 SerLeuAspGluAlaTyrArgPheValLysGluLysArgProThrIleSerProAsn 137  
 Db 3 TCTTAGATGAAGCTTACAGATTGTGTAAGAAAAAGACCTACTATATCTCAACTTC 62  
 Qy 138 AsnPheLeuGlnLeuLeuAspTyrGluLysLysIleLysAsnGlnThrGlyAlaSer 157



QY 317 aSerIleProLysLeuGlnThrAlaArgProSerAspSerGlnSerLysArgLeuHi 337  
 Db 602 CAGCATCCCCAAGAAGCTGCAGACCCCGAGGCTTCAGACGACGAGCAAGCGATTGCA 661  
 QY 337 sSerValArgThrSerSerGlyThrAlaGlnArgSerLeuLeuSerProLeuHisAr 357  
 Db 662 TTCGGTCAGAACACGACGACGAGTGGCAGCCGCCACGAGGTCCCTTTATCTCCACTGCATCG 721  
 QY 357 gSerGlySerValGluAspAsnTyrHisThrSerPheLeu-PheGlyLeuSerThrSerG 377  
 Db 722 AAGTGGGAGCGTGGAGACAAATACCACACACGAGCTTCCTTTTCGGCCCTTCCACACGCC 781  
 QY 377 lnglnHisLeuThrLysSerAlaGlyLeuGly 387  
 Db 782 AGCAGCACTTCAGAGTCTGCTGGCTGGGG 813

RESULT 7  
 BQ951695  
 LOCUS AGENCOURT 8947130 NIH\_MGC\_101 Homo sapiens cDNA clone IMAGE:6461668  
 DEFINITION 5', mRNA sequence.

ACCESSION BQ951695  
 VERSION BQ951695.1 GI:22367173  
 KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 974)

AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

CONTACT: Robert Strausberg, Ph.D.

EMAIL: cga@bbs-remail.nih.gov

Tissue Procurement: ATCC

CNA Library Preparation: Rubin Laboratory

CNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Information found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LICM2650 row: f column: 05

High quality sequence stop: 617.

## FEATURES

source

1..974

Location/Qualifiers

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:6461668"

/tissue\_type="epidermoid carcinoma, cell line"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="NIH\_MGC\_101"

/note="Organ: lung; Vector: pOTB7; Site 1: EcoRI; Site 2:

XhoI; cDNA made by oligo-dt priming. Directionally cloned

into EcoRI/XhoI sites using the following 5' adaptor:

GGCAGCAG(G). Library constructed by Ling Hong in the

laboratory of Gerald M. Rubin (University of California,

Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and

Superscript II RT (Life Technologies). Note: this is a

NIH\_MGC Library."

## ORIGIN

Alignment Scores:  
 Pred. No.: 3,36e-138 Length: 974  
 Score: 160.00 Matches: 241  
 Percent Similarity: 99.18% Conservative: 0  
 Best Local Similarity: 99.18% Mismatches: 1  
 Query Match: 30.95% Indels: 2  
 Ds: 13 Gaps: 0

US-09-964-277-21 (1-517) x BQ951695 (1-974)

QY 118 SerIleuAspGluAlaTyrArgPheValLysGluLysArgProThrIleSerProAsnPhE 137

Db 3 TCTTTAGATGAGCTTACAGATTGTGTAAGAAAAGAACCTACTATATCTCCAACTTC 62  
 QY 138 AsnPheLeuGlyGlnLeuLeuAspTyrGlnLysLysIleLysAsnGlnThrGlyAlaSer 157  
 Db 63 AATTTTCTGGCCAACTCTCTGACTATGAGAGAAGATTAAAGAACCAGACTGGAGCATCA 122  
 QY 158 GlyProLysSerLysLeuLeuLeuHisLeuGluLysProAsnGluProValProAla 177  
 Db 123 GGGCCAAAGAGCAAACTCAGCTGCTGCACCTGGAGAAACCAATGAACCTGTCCTGCT 182  
 QY 178 ValSerGluGlyGlyGlnLysSerGlnThrProLeuSerProProCysAlaAspSerAla 197  
 Db 183 GTCTCAGAGGGTGGACAGAAAAGCAGACGCGCTCAGTCCACCTGTGCGACTGCT 242  
 QY 198 ThrSerGluAlaAlaGlyGlnArgProValHisProAlaSerValProSerValProSer 217  
 Db 243 ACCTCAGAGCGCAGGACAAAGGCCCTGTGCATCCCGCCAGCGTCCCGGCTGCCAGC 302  
 QY 218 ValGlnProSerLeuLeuGluAspSerProLeuValGlnAlaLeuSerGlyLeuHisLeu 237  
 Db 303 GTGCAGCGCTCGCTGTAGAGGACAGCGCGCTGTACAGGGCTCAGTGGGCTGCACCTG 362  
 QY 238 SerAlaAspArgLeuGluAspSerAsnLysLysLysArgSerPheSerLeuAspIleLys 257  
 Db 363 TCCGCAGACAGCGTGGAGACAGCAATAAGCTCAAGCGTCTCTCTCTGGATATCAA 422  
 QY 258 SerValSerTyrSerAlaSerMetAlaAlaSerLeuHisGlyPheSerSerSerGluAsp 277  
 Db 423 TCAGTTTCATATTCAGCAGCATGGCAGCATCTTACATGGCTTCTCTCATCAGAAGAT 482  
 QY 278 Ala-LeuGluTyrTyrLysProSerThrThrLeuAspGlyThrAsnLysLeuCysGlnPh 297  
 Db 483 GT-TTTGGAATACTACAAACCTTCCACTACTCTGGATGGACCAACAGCTATGCCAGTT 541  
 QY 297 eSerProValGlnGluLeuSerGluGlnThrProGluThrSerProAspLysGluGluAl 317  
 Db 542 CTCCTCTGTTCAGAACTATCGAGCAGACTCCCGAAACCACTCTCTGATAGGAGGAGC 601  
 QY 317 aSerIleProLysLysLeuGlnThrAlaArgProSerAspSerGlnSerLysArgLeuHi 337  
 Db 602 CAGCATCCCCAAGAAGCTGCAGACCGCCAGGCTTCAGACAGCCAGAGCAAGCGATTGCA 661  
 QY 337 sSerValArgThrSerSerGlyThrAlaGlnArgSerLeuLeuSerProLeuHisAr 357  
 Db 662 TTCGGTCAGAACACGACGAGTGGCAGCCGCCAGAGGTCCTTTATCTCCACTGCATCG 721  
 QY 357 gSerGly 359  
 Db 722 AAGTGGG 728

## RESULT 8

BX479029

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BX479029 663 bp mRNA linear EST 04-SEP-2003  
 DXFZP5686722208\_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone  
 DXFZP5686722208\_5', mRNA sequence.

BX479029

BX479029.1 GI:31914415

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 663)

Bahr,A., Lauber,J., Mewes,H.W., Weil,B., Amid,C., Osanger,A.,

Foto,G., Han,M. and Wiemann,S.

EST (Bahr,A., Lauber,J., Mewes,H.W., Weil,B., et al.)

Unpublished (2003)

Contact: MIPS

MIPS

Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany

This is the 5' sequence of the clone insert

Clone from S. Wiemann, Molecular Genome Analysis, German Cancer



Research Center (DKFZ); Email s.wiemann@kfz-heidelberg.de;  
sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing  
consortium of the German Genome Project.

No s1 sequence available.  
This clone (DKFZp686J22208) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

# FEATURES

source  
1..563  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="DKFZp686J22208"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/clone\_lib="686 (synonym: hlcc3)"  
/note="Vector: pTriplex2; Site\_1: SfiIA; Site\_2: SfiIB;  
cDNA-collection"

## ORIGIN

Alignment Scores:  
Pred. No.: 1.93e-137 Length: 663  
Score: 159.00 Matches: 190  
Percent Similarity: 98.96% Conservative: 0  
Best Local Similarity: 98.96% Mismatches: 1  
Query Match: 30.75% Indels: 2  
DB: 13 Gaps: 0

US-09-964-277-21 (1-517) x BX479029 (1-663)

QY 92 ValLeuValHisCysLeuAlaGlySerArgSerAlaThrIleAlaLeuValrille 111  
DB 57 GTTCTAGTGCACTTTAGCTGGGATCTCCGCTCGCCACCATGCTATCGCTACATC 116  
QY 112 MetLysArgMetAspMetSerLeuAspGluAlaTyrArgPheValLysGluLysArgPro 131  
DB 117 ATGAGAGGATGGACATGCTTTAGATGAAGCTTACAGATTGTGAAGAAAAAGACCT 176  
QY 132 ThrIleSerProAanPheAanPheLeuGlyGlnLeuLeuAspTyrGluLysLysIleLys 151  
DB 177 ACTATATCTCCAAACTTCAATTTCTGGGCCAACTCTCGACTATGAGAAGATTAG 236  
QY 152 AsnGlnThrGlyAlaSerGlyProLysSerLysLysLysLeuHisLeuGluLysPro 171  
DB 237 AACCAAGCTGGAGCATCAGGCCCAAGAGCAAACTCAAGCTGTGCACTGGAGAGCCA 296  
QY 172 AsnGluProValProAlaValSerGluGlyGlyGlnLysSerGluThrProLeuSerPro 191  
DB 237 AATGAACTGTCTCTGTCTGCTCAGAGGGTGGACAGAAAGCGAGCGCCCTCAGTCCA 356  
QY 192 ProCysAlaAspSerAlaThrSerGluAlaAlaGlyGlnArgProValHisProAlaSer 211  
DB 357 CCCTGTGCGGACTCTGTACTCTCAGAGCGCAGCAGCAAAAGGCCGCTGCATCCCGCAGC 416  
QY 212 ValProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuValGlnAla 231  
DB 417 GTGCCACGGTGGCCAGCGTGCAGCGCTGCTGTGTAGAGGACGCCGCTGGTACAGGGC 476  
QY 232 LeuSerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnLysLysLysArgSer 251  
DB 477 CTCAGTGGGCTGCACCTGTCCGACAGCAGCGTGGAGACAGCAATAAGCTCAAGCGTNC- 535  
QY 252 -PheSerLeuAspLysSerValSerTyrSerAlaSerMetAlaAlaSerLeuHisG1 271  
DB 536 CTTCTCTGTGATATCAAAATCAGTTTCATATTCAGCCAGCATGCGCATCTTACATGG 595  
QY 271 yPheSerSerSerGluAspAlaLeuGluTyrTyr 282  
DB 596 CTTCTCTCTATCAGAGATGCTTTGGAACTACTAC 629

RESULT 9  
BQ945892  
LOCUS BQ945892 883 bp mRNA linear EST 21-AUG-2002

DEFINITION AGENCOURT\_8926314 NIH\_MGC\_101 Homo sapiens cDNA clone IMAGE:6463045  
5', mRNA sequence.  
ACCESSION BQ945892  
VERSION BQ945892.1 GI:22361370  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
NIH-MGC http://mgc.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: gspbs-remail.nih.gov  
Tissue Procurement: AFCC  
CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LICM2653 row: 0 column: 14  
High quality sequence stop: 672.

## FEATURES

source  
1..883  
/location="Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6463045"  
/tissue\_type="epidermoid carcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH MGC 101"  
/note="Organ: lung; Vector: pOTB7; Site 1: EcoRI; Site 2:  
XhoI; cDNA made by oligo-dT priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:  
GGCACCAG(G). Library constructed by Ling Hong in the  
Laboratory of Gerald M. Rubin (University of California,  
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
Superscript II RT (Life Technologies). Note: this is a  
NIH\_MGC Library."

## ORIGIN

Alignment Scores:  
Pred. No.: 1.14e-118 Length: 883  
Score: 139.00 Matches: 249  
Percent Similarity: 98.42% Conservative: 0  
Best Local Similarity: 98.42% Mismatches: 2  
Query Match: 26.89% Indels: 4  
DB: 13 Gaps: 0

US-09-964-277-21 (1-517) x BQ945892 (1-883)

QY 139 PheLeuGlyGlnLeuLeuAspTyrGluLysLysLysLysAsnGlnThrGlyAlaSerGly 158  
DB 63 TTCTGGGCCCACTCTCTGGACTATGAGAAGAGATTAGAACACAGCTGGACATCAGGG 122  
QY 159 ProLysSerLysLysLeuHisLeuGluLysProAsnGluProValProAlaVal 178  
DB 123 CCAAGAGCAAACTCAAGCTGTGCACTGGAGAGCCAAATGAACCTGTCCCTGCTGTC 182  
QY 179 SerGluGlyGlyGlnLysSerGluThrProLeuSerProCysAlaAspSerAlaThr 198  
DB 183 TCAGAGGGTGGACAGAAAGCGAGAGCGCCCTCAGTCCACCTGTGCCACTCTGTACC 242  
QY 199 SerGluAlaAlaGlyGlnArgProValHisProAlaSerValProSerValProSerVal 218  
DB 243 TCAGAGGACAGCAGCAAAAGGCCGCTGCATCCGCCAGCGTGCACCGAGCGTG 302  
QY 219 GlnProSerLeuLeuGluAspSerProLeuValGlnAlaLeuSerGlyLeuHisLeuSer 238  
DB 303 CAGCCGCTGCTGTAGAGGACAGCCGCTGGTACAGCGCTCAGTGGGCTGACCTGTC 362



/tissue\_type="mixed"  
/dev\_stage="mixed"  
/lab\_host="DH10B (Life Technologies)"  
/clone\_lib="NCI CGAP Sub9"  
/notes="Vector: p773-Pac (Pharmacia) with a modified  
polylinker; Site 1: EcoR I; Site 2: Not I; tissues:  
Cholonic mucosa with Crohn's disease, Cholonic mucosa with  
ulcerative colitis, Fetal thymus, Cervix, Cervical  
adenocarcinoma, Bladder carcinoma, Ligament cells, Prostate  
carcinoma, Brain oligodendrocyte  
NCI CGAP Sub9 is a subtracted cDNA library constructed  
according to Bonaldo, Lennon and Soares, Genome Research,  
6:791-806, 1996. First strand cDNA synthesis was primed  
with an oligo-dT primer containing a Not I site. Double  
stranded cDNA was ligated to an EcoR I adaptor, digested  
with Not I, and cloned directionally into p773-Pac  
vector. The oligonucleotide used to prime the synthesis of  
first-strand cDNA contains a library tag sequence that is  
located between the Not I site and the (dT)18 tail. The  
sequence tags for this library are CGTC, AACG, GGCGC,  
GGAG, TAGC, TAGC, ATGG, AGCA, ATCAC. For additional  
information, contact: Bento Soares, bento-soares@uiowa.edu  
TAG\_TISSUE=Prostate Carcinoma  
TAG\_LIB=UI-H-CO  
TAG\_SEQ=ATGG"

## ORIGIN

Alignment Scores:  
Pred. No.: 7,33e-107 Length: 444  
Score: 126.00 Matches: 126  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 24.37% Indels: 0  
DB: 12 Gaps: 0

US-09-964-277-21 (1-517) x BM987142 (1-444)

Qy 392 HisSerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTrpTyrPhe 411  
Db 443 CACTCGGATATCTGGCCCCCAGAGCTTACCCCTTCCCTGACCAGCAGCTGGTATTT 384  
Qy 412 AlaThrGluSerSerHisPheTyrSerAlaSerAlaIleTyrGlySerAlaSerTyr 431  
Db 383 GCCACAGAGCTCTACACTTCTGCTCAGCCACTACGGAGGCGAGTCCAGTTAC 324  
Qy 432 SerAlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerValArg 451  
Db 323 TCTGCTTACAGCTGACGAGCTGCCACTTGGGAGACCAAGTCTATTCTGTGGCAGG 264  
Qy 452 ArgGlnLysProSerAspAlaAspSerArgArgSerTrpHisGluSerProPhe 471  
Db 263 CGGCAGAAAGCCAGTGACAGAGCTGACTCGCGCGAGCTGGCATGAAGAGAGCCCTTT 204  
Qy 472 GluLysGlnPheLysArgSerCysGlnMetGluPheGlyGluSerIleMetSerGlu 491  
Db 203 GAAAGCAGTTTAAACGACAGAGCTGCCAATGGAATTTGGAGAGAGCATATGTCAGAG 144  
Qy 492 AsnArgSerArgGluGluLeuGlyLysValGlySerGlnSerSerPheSerGlySerMet 511  
Db 143 AACAGGTCACGGGAAGAGCTGGGMAAGTGGCGAGTCAGTCTAGCTTTTCGGCGCAGCATG 84  
Qy 512 GluIleIleGluValSer 517  
Db 83 GAAATCATTTGAGGTCTCC 66

## RESULT 12

BE897795  
LOCUS 60143457F1 NIH\_MGC\_72 Homo sapiens cDNA clone IMAGE:3923247 5',  
DEFINITION mRNA sequence.  
ACCESSION BE897795  
VERSION BE897795.1 GI:10363618  
KEYWORDS EST.

SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 920)  
AUTHORS NIH-MGC http://img.nci.nih.gov/  
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabbs@mail.nih.gov  
Tissue Procurement: ATCC/DCTD/DTP  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM9759 row: b column: 16  
High quality sequence stop: 678.

FEATURES  
Location/Qualifiers  
1..920  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3923247"  
/tissue\_type="melanotic melanoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH MGC 72"  
/notes="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 2 Kb. Library constructed by Life  
Technologies."

## ORIGIN

Alignment Scores:  
Pred. No.: 1,56e-106 Length: 920  
Score: 126.00 Matches: 126  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 24.37% Indels: 0  
DB: 10 Gaps: 0

US-09-964-277-21 (1-517) x BE897795 (1-920)

Qy 90 GlyCysValLeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAla 109  
Db 18 GGATGTGTTCTAGTGACATGTTTAGTGGGATCTCCGCTCCGCCACCATCGCTATCGCC 77  
Qy 110 TyrIleMetLysArgMetAspMetSerLeuAspGluAlaTyrArgPheValLysGluLys 129  
Db 78 TACATCATGAAGAGGATGGACATGCTTTTAGATGAAGCTTACAGATTGTGAAGAAAA 137  
Qy 130 ArgProThrIleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGluLysLys 149  
Db 138 AGACCTACTATATCTCAAACTTCAATTTTCTGGCCAACTCTCTGGACTATGAGAAGAAG 197  
Qy 150 IleLysAsnGlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuHisLeuGlu 169  
Db 198 ATTAAGAACCCAGACTGGAGCATCAGGGCCAAAGAGCAAACTCAAGCTGTCACCTGGAG 257  
Qy 170 LysProAsnGluProValProAlaValSerGluGlyGlyGlnLysSerGluThrProLeu 189  
Db 258 AAGCCAAATGAACCTGCTCTGCTCTCAGAGGGTGGACAGAAAAGCAGACGCCCTC 317  
Qy 190 SerProProCysAlaAspSerAlaThrSerGluAlaAlaGlyGlnArgProValHisPro 209  
Db 318 AGTCCACCTGTGCTCCACTCTGCTACTCAGAGGCGAGGACAGGCAAAAGGCCGTCATCCC 377  
Qy 210 AlaSerValProSerVal 215  
Db 378 GCCAGGTGCCCGCAGCGTG 395

## RESULT 13

AV703072  
 LOCUS AV703072 621 bp mRNA linear EST 08-OCT-2000  
 DEFINITION AV703072 ADB Homo sapiens cDNA clone ADBDAB06 5', mRNA sequence.  
 ACCESSION AV703072  
 VERSION AV703072.1 GI:10719402  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 621)  
 AUTHORS Peng, Y., Song, H., Huang, Q., Huang, C., Gu, Y., Yang, Y., Gao, G.,  
 Xiao, H., Xu, X., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z.,  
 Xu, Z., Zeng, L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S.,  
 Zhong, M., Lu, G., Hu, R., Chen, J., Chen, Z. and Han, Z.  
 TITLE Homo sapiens cDNA ADB clones  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Zeguang Han  
 Chinese National Human Genome Center at Shanghai  
 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai  
 201203, P. R. China  
 Tel: 86-21-50801919(ex.45)  
 Fax: 86-21-50801922  
 Email: hanzg@chgc.sh.cn  
 This clone is available at CHGC in Shanghai.  
 FEATURES  
 source  
 Location/Qualifiers  
 1..621  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="ADB06"  
 /tissue\_type="Adrenal gland"  
 /dev\_stage="Adult"  
 /lab\_host="SOLR"  
 /clone\_lib="ADB"  
 /note="Vector: pBluescript sk(-); Site\_1: EcoRI; Site\_2:  
 XhoI"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 7,59e-105 Length: 621  
 Score: 124.00 Matches: 124  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 23.98% Indels: 0  
 DB: 9 Gaps: 0  
 US-09-964-277-21 (1-517) x AV703072 (1-621)  
 QY 202 AlaGlyGlnArgProValHisProAlaSerValProSerValProSerValGlnProSer 221  
 Db 138 GCAGGACAAAGGCCCGTGCATCCGCCAGCGTGCAGCGTGCAGCGTGCAGCGTGC 197  
 QY 222 LeuGluAspSerProLeuValGlnAlaLeuSerGlyLeuHisLeuSerAlaAspArg 241  
 Db 198 CTGTTAGAGACAGCCCGTGTGTACAGGCGCTCATGGGCTGCACCTGTCCGACAGG 257  
 QY 242 LeuGluAspSerAsnLysLeuLysArgSerPheSerLeuAspIleLysSerValSerTyr 261  
 Db 258 CTGGAAGACAGCAATAAGCTCAAGCGTCTCTCTCTCTCGATATCAATCAGTTTCATAT 317  
 QY 262 SerAlaSerMetAlaLeuSerHisGlyPheSerSerGluAspAlaLeuGluTyr 281  
 Db 318 TCAGCAGCAGTGCAGCATCTTACATGCTTCTCTCTCTATCAGAAGATGCTTTGAATAC 377  
 QY 282 TyrLysProSerThrThrLeuAspGlyThrAsnLysLeuCysGlnPheSerProValGln 301  
 Db 378 TACAACTTCCTCACTACTCTGGATGGACCAACAAGCTATGCCAGTCTTCCCTGTTTCAG 437  
 QY 302 GluLeuSerGluGlnThrProGluThrSerProAspLysGluGluAlaSerIleProLys 321  
 Db 438 GAACTATCGAGAGAGATCCCGAAACAGTCTCTGTATAGGAGAGAGCCAGCATCCCCAAG 497

QY 322 LysLeuGlnThr 325  
 Db 498 AAGTGCAGACC 509  
 RESULT 14  
 LOCUS AV701628  
 DEFINITION AV701628 ADB Homo sapiens cDNA clone ADBCHE02 5', mRNA sequence.  
 ACCESSION AV701628  
 VERSION AV701628.1 GI:10717958  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 693)  
 AUTHORS Peng, Y., Song, H., Huang, Q., Huang, C., Gu, Y., Yang, Y., Gao, G.,  
 Xiao, H., Xu, X., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z.,  
 Xu, Z., Zeng, L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S.,  
 Zhong, M., Lu, G., Hu, R., Chen, J., Chen, Z. and Han, Z.  
 TITLE Homo sapiens cDNA ADB clones  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Zeguang Han  
 Chinese National Human Genome Center at Shanghai  
 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai  
 201203, P. R. China  
 Tel: 86-21-50801919(ex.45)  
 Fax: 86-21-50801922  
 Email: hanzg@chgc.sh.cn  
 This clone is available at CHGC in Shanghai.  
 FEATURES  
 source  
 Location/Qualifiers  
 1..693  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="ABCHE02"  
 /tissue\_type="Adrenal gland"  
 /dev\_stage="Adult"  
 /lab\_host="SOLR"  
 /clone\_lib="ADB"  
 /note="Vector: pBluescript sk(-); Site\_1: EcoRI; Site\_2:  
 XhoI"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 8.5e-105 Length: 693  
 Score: 124.00 Matches: 124  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 23.98% Indels: 0  
 DB: 9 Gaps: 0  
 US-09-964-277-21 (1-517) x AV701628 (1-693)  
 QY 202 AlaGlyGlnArgProValHisProAlaSerValProSerValProSerValGlnProSer 221  
 Db 138 GCAGGACAAAGGCCCGTGCATCCGCCAGCGTGCAGCGTGCAGCGTGCAGCGTGC 197  
 QY 222 LeuGluAspSerProLeuValGlnAlaLeuSerGlyLeuHisLeuSerAlaAspArg 241  
 Db 198 CTGTTAGAGACAGCCCGTGTGTACAGGCGTGCAGCGTGCAGCGTGCAGCGTGCAGG 257  
 QY 242 LeuGluAspSerAsnLysLeuLysArgSerPheSerLeuAspIleLysSerValSerTyr 261  
 Db 258 CTGGAAGACAGCAATAAGCTCAAGCGTCTCTCTCTCTCTGATATCAATCAGTTTCATAT 317  
 QY 262 SerAlaSerMetAlaLeuSerHisGlyPheSerSerGluAspAlaLeuGluTyr 281  
 Db 318 TCAGCAGCAGTGCAGCATCTTACATGCTTCTCTCTCTATCAGAAGATGCTTTGAATAC 377  
 QY 282 TyrLysProSerThrThrLeuAspGlyThrAsnLysLeuCysGlnPheSerProValGln 301  
 Db 378 TACAACTTCCTCACTACTCTGGATGGACCAACAAGCTATGCCAGTCTTCCCTGTTTCAG 437

QY 302 GluLeuSerGluInThrProGluThrSerProAspLysGluGluAlaSerIleProLys 321  
 Db 438 GAACTATCGGAGCAGACTCCCGAACAACAGTCTGATAGGAGGAGGAGCCAGCATCCCCAAAG 497  
 QY 322 LysLeuGlnThr 325  
 Db 498 AAGCTGCAGACC 509

RESULT 15  
 B1917706/c  
 LOCUS B1917706 660 bp mRNA linear EST 16-OCT-2001  
 DEFINITION 603183686P1 NIH\_MGC\_121 Homo sapiens cDNA clone IMAGE:5247884 5',  
 mRNA sequence.  
 ACCESSION B1917706  
 VERSION B1917706.1 GI:16181504  
 SOURCE EST.  
 ORGANISM Homo sapiens (human)

REFERENCE 1 (bases 1 to 660)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 NIH-MGC <http://mgs.nci.nih.gov/>.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)

COMMENT  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: Life Technologies, Inc.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LLAM11625 row: k column: 21  
 High quality sequence stop: 660.

FEATURES  
 source  
 1..660  
 Location/Qualifiers  
 1..660  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clones="IMAGE:5247884"  
 /lab\_host="DH10B"  
 /clone\_lib="NIH\_MGC\_121"  
 /note="Organ: Brain; Vector: pCMV-SPORT6; Site 1: NotI;  
 Site 2: EcoRV (destroyed); RNA source anonymous pool of 3  
 fetal brains, female age 20 weeks, female age 24 weeks,  
 and male age 26 weeks. Library is oligo-dT primed and  
 directionally cloned (EcoRV site is destroyed upon  
 cloning). Average insert size 1.7 kb, insert size range  
 0.7-3.5 kb. Library is normalized and enriched for  
 full-length clones and was constructed by C. Gruber  
 (Invitrogen). Research Genetics tracking code 017. Note:  
 this is a NIH\_MGC Library."

## ORIGIN

Alignment Scores:  
 Pred. No.: 2,22e-83 Length: 660  
 Score: 101.00 Matches: 127  
 Percent Similarity: 98.45% Conservative: 0  
 Best Local Similarity: 98.45% Mismatches: 0  
 Query Match: 19.54% Indels: 2  
 DB: 12 Gaps: 0

US-09-964-277-21 (1-517) x B1917706 (1-660)

QY 89 AsnGlyCysValLeuValHisCysLeuAlaGlyIleSerArgSer-AlaThrIleAlaIrl 108  
 Db 658 AATGGATGTGTTCTAGTGCATGTTTACTGGGATCTCCCGCTCCGACCATTCGCTAT 599  
 QY 108 ealaTyrlleMetLysArgMetAspMetSerLeuAspGluAlaTyArgPheValLysGI 128  
 Db 598 CGCCTACATCATGAAGAGGATGGACATGCTTTAGATGAAGCTTACAGATTTGTGAAGA 539

QY 128 uLysArgProThrIleSerPro-AsnPheAsnPheLeuGlyGlnLeuLeuAspTyArgLUL 148  
 Db 538 AAAAAGACCTACTATATCTCCAGAACTTCAATTTTCTGGGCCCAACTCTCTGGACTATGAGA 479  
 QY 148 yLysIleLeuLysAsnGlnThrGlyValAsnGlyProLysSerLysLeuLysLeuHisl 168  
 Db 478 AGAAGATTAAAGAACAGACTGGAGCATCAGGGCCAAAGAGCAAACTCAAGCTGCTGCACC 419  
 QY 168 euGluLysProAsnGluProValProAlaValSerGluGlyGlnLysSerGluThrP 188  
 Db 418 TGGAGAAAGCCAAATGAACCTGTCTCCCTGCTCTCAGAGGGTGGACAGAAAAGCGAGACGC 359  
 QY 188 rLysSerProProCysAlaAspSerAlaThrSerGluAlaAlaGlyGlnArgProValH 208  
 Db 358 CCCTCAGTCCACCTGTGCGGACTCTGCTACCTCAGAGGCGAGCAAGAGCCCGTGC 299  
 QY 208 sProAlaSerValProSerVal 215  
 Db 298 ATCCCGCCAGCGTACCCAGCGTG 276

Search completed: February 28, 2004, 04:12:25  
 Job time : 3862 secs

GenCore version 5.1.6  
Copyright (C) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 27, 2004, 18:39:38 ; Search time 12757 Seconds  
(without alignments)  
11320.764 Million cell updates/sec

Title: US-09-964-277-20  
Perfect score: 3332  
Sequence: 1 gagagaaggagaataaata.....ataaagatgaacttggtttc 3332

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_htg.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mu.\*
- 20: em\_om.\*
- 21: em\_or.\*
- 22: em\_ov.\*
- 23: em\_pat.\*
- 24: em\_ph.\*
- 25: em\_pl.\*
- 26: em\_ro.\*
- 27: em\_sts.\*
- 28: em\_un.\*
- 29: em\_vi.\*
- 30: em\_htg\_hum.\*
- 31: em\_htg\_inv.\*
- 32: em\_htg\_other.\*
- 33: em\_htg\_mus.\*
- 34: em\_htg\_pln.\*
- 35: em\_htg\_rod.\*
- 36: em\_htg\_mam.\*
- 37: em\_htg\_vrt.\*
- 38: em\_sy.\*
- 39: em\_htgo\_hum.\*
- 40: em\_htgo\_mus.\*
- 41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	3332	100.0	3332	6	AX441229	AX441229 Sequence
2	3158	94.8	3496	6	AX441210	AX441210 Sequence
3	3145	94.4	3566	9	AF506796	AF506796 Homo sapi
4	3133.4	94.0	3521	9	AB052156	AB052156 Homo sapi
5	3118.8	93.6	3766	6	AX374994	AX374994 Sequence
6	3106.2	93.2	5450	6	AX482439	AX482439 Sequence
7	3106.2	93.2	5450	6	AX482478	AX482478 Sequence
8	2950	88.5	3544	6	AX260340	AX260340 Sequence
9	2924.6	87.8	5111	6	AX482372	AX482372 Sequence
10	2888	86.7	3284	9	BC042101	BC042101 Homo sapi
11	2766.4	83.0	4790	6	BD171157	BD171157 Novel gen
12	2766.4	83.0	4790	6	BD183422	BD183422 Novel gen
13	2766.4	83.0	4790	6	AB051487	AB051487 Homo sapi
14	2717.2	81.5	3104	6	AX405700	AX405700 Sequence
15	2710	81.3	3059	6	AX278461	AX278461 Sequence
16	2437.4	73.2	2807	9	BC031643	BC031643 Homo sapi
17	2390.8	71.8	2732	6	AX180875	AX180875 Sequence
18	2111.4	63.4	172206	9	AC007619	AC007619 Homo sapi
19	2111.4	63.4	183344	2	AC131617	AC131617 Homo sapi
20	1936	58.1	4827	10	AF345952	AF345952 Mus muscu
21	1935	58.1	1935	9	AY038927	AY038927 Homo sapi
22	1822.6	54.7	4943	10	BC059232	BC059232 Mus muscu
23	1822.6	54.7	4975	10	BC057321	BC057321 Mus muscu
24	1818	54.6	4874	10	AB052157	AB052157 Mus muscu
25	1795.4	53.9	1916	6	AX835305	AX835305 Sequence
26	1795.4	53.9	1916	9	AK098310	AK098310 Homo sapi
27	1774	53.2	4932	10	AF345951	AF345951 Mus muscu
28	1762.4	52.9	2102	6	AX713989	AX713989 Sequence
29	1762.4	52.9	2102	9	AK055973	AK055973 Homo sapi
30	1742	52.3	2200	6	AX921915	AX921915 Sequence
31	1741	52.3	2118	6	AX099933	AX099933 Sequence
32	1712.4	51.4	2071	6	AX921917	AX921917 Sequence
33	1660	49.8	1998	6	AX260342	AX260342 Sequence
34	1638.2	49.2	2756	6	AX482444	AX482444 Sequence
35	1133	34.0	201474	2	AC126692	AC126692 Mus muscu
36	1133	34.0	236589	2	AC118035	AC118035 Mus muscu
37	1133	34.0	250782	2	AC128093	AC128093 Rattus no
38	964.2	28.9	242590	2	AC133722	AC133722 Rattus no
39	962.4	28.9	244605	2	AC097818	AC097818 Rattus no
40	886.8	26.6	2558	6	AX748225	AX748225 Sequence
41	886.8	26.6	2558	9	AK093617	AK093617 Homo sapi
42	841.4	25.3	3861	10	AF345954	AF345954 Mus muscu
43	679.4	20.4	4026	10	AF345953	AF345953 Mus muscu
44	595	17.9	787	6	BD019102	BD019102 Novel gen
45	595	17.9	787	6	BD083796	BD083796 Nucleic a

ALIGNMENTS

RESULT 1	AX441229	AX441229	3332 bp	DNA	linear	PAT 28-JUN-2002
LOCUS	AX441229	Sequence 20 from Patent WO0226597.				
DEFINITION	AX441229					
ACCESSION	AX441229					
VERSION	AX441229.1	GI:21665771				
KEYWORDS						
SOURCE 1						
ORGANISM	Homo sapiens (human)					
	Homo sapiens					
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;					
	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.					
REFERENCE						
AUTHORS	Luche, R.M. and Wei, B.					
TITLE	Dsp-16 dual-specificity phosphatase					
JOURNAL	Patent: WO 0226597-A 20 04-APR-2002;					

[illegible]







[illegible][illegible]

Db 373 GTCACTCTTTATGAATCAAAATGTCAGGGCTGCTTTTGTGACGGAGTCTTTTTCGAAGAG 432  
Qy 362 CACATCAACGGGAAGAGAGAGACATTCACCTTGGAGGGCTCTTGTGAAATGGSTTT 421  
Db 433 CACATCAACGGGAAGAGAGAGACATTCACCTTGGAGGGCTCTTGTGAAATGGSTTT 492  
Qy 422 AACTCTCCCTTTTGCAGTCAACACAGAGCTGACCTCATACACTTTTATAGTACAATGGAGTG 481  
Db 493 AACTCTCCCTTTTGCAGTCAACACAGAGCTGACCTCATACACTTTTATAGTACAATGGAGTG 552  
Qy 482 GCTGAGCCTTTGAGACACACACATTAATCATCATCTGTCGCAAAATTAAGAAGAGGTGGGA 541  
Db 553 GCTGAGCCTTTGAGACACACACATTAATCATCATCTGTCGCAAAATTAAGAAGAGGTGGGA 612  
Qy 542 AAAGAGGACTTATGTTGTCATGGCCCATGAGATGANTTGAAGTCAAAATGTTACTGAGA 601  
Db 613 AAAGAGGACTTATGTTGTCATGGCCCATGAGATGANTTGAAGTCAAAATGTTACTGAGA 672  
Qy 602 GGTGCTGCTCTGCTGGAAGTGAACGGAAGAAAGTCTGCTAAATGATAGCCGGCCAT 661  
Db 673 GGTGCTGCTCTGCTGGAAGTGAACGGAAGAAAGTCTGCTAAATGATAGCCGGCCAT 732  
Qy 662 TTGTGGAATACAAATACATCCACATTTTGGAGCCATTAATCAACTGCTCCAGCTTA 721  
Db 733 TTGTGGAATACAAATACATCCACATTTTGGAGCCATTAATCAACTGCTCCAGCTTA 792  
Qy 722 TGAAGCGAAGTTGCAACAGGACAAAGTGTAAATACAGAGCTCATCCAGCATTCAGCGA 781  
Db 793 TGAAGCGAAGTTGCAACAGGACAAAGTGTAAATACAGAGCTCATCCAGCATTCAGCGA 852  
Qy 782 AACATAGGTTGACATGATTCAGTCAAGAGTTGATGTTTACGATCAAGGCTCCCAAG 841  
Db 853 AACATAGGTTGACATGATTCAGTCAAGAGTTGATGTTTACGATCAAGGCTCCCAAG 912  
Qy 842 ATGTTGCTCTCTCTCTCTCAGAGTGTCTTCTCACTGTACTTCTGGGTAACTGGAGAGA 901  
Db 913 ATGTTGCTCTCTCTCTCTCAGAGTGTCTTCTCACTGTACTTCTGGGTAACTGGAGAGA 972  
Qy 902 GCTTCAACTCTGTTCACTGCTTGC----- 926  
Db 973 GCTTCAACTCTGTTCACTGCTTGCAGTGGGTTTGTGAGTCTCTGCTGTTTTCCTG 1032  
Qy 927 ----- 926  
Db 1033 GCCTCTGTGAGGAAATCCACTCTAGTCCCTACCTGCATTTCTCAGGCTTGTCTTACCTG 1092  
Qy 927 ----- 926  
Db 1093 TTGCCAACAATTGGGCCAACCCGAATCTTCCCAATCTTTATCTTGGCTGCCAGGAGATG 1152  
Qy 927 -----AGGAGCTGATGCACAGAAATGGGATTTGTTATGTTAAATGCCAGCATTA 977  
Db 1153 TCCTCAACAGGAGCTGATGCAGAAATGGGATTTGTTATGTTAAATGCCAGCATTA 1212  
Qy 978 CCTGTCCAAAGCCTGACTTTATCCCGAGTCTCAITTCCTGCGTGTGCTGTGAATGACA 1037  
Db 1213 CCTGTCCAAAGCCTGACTTTATCCCGAGTCTCAITTCCTGCGTGTGCTGTGAATGACA 1272  
Qy 1038 GCTTTGTGAGAAATTTTGCCTGTGTGGAACAAATCAGTAGATTTCAITGAGAAAGCAA 1097  
Db 1273 GCTTTGTGAGAAATTTTGCCTGTGTGGAACAAATCAGTAGATTTCAITGAGAAAGCAA 1332  
Qy 1098 AAGCTCCCAATGATGTTCTAGTGCATCTGTTAGCTGGGATCTCCGCTCCGCCACCA 1157  
Db 1333 AAGCTCCCAATGATGTTCTAGTGCATCTGTTAGCTGGGATCTCCGCTCCGCCACCA 1392  
Qy 1158 TCGTATTCGCTCATCATGAAGAGATGGACATGTCCTTTAGATGAAGCTTACAGTTTG 1217  
Db 1393 TCGTATTCGCTCATCATGAAGAGATGGACATGTCCTTTAGATGAAGCTTACAGTTTG 1452  
Qy 1218 TGAAGAAABAAGACCTTACTATATCTCCAAACTTCAATTTCTGGGCCAACTCTCTGGACT 1277  
Db 1453 TGAAGAAABAAGACCTTACTATATCTCCAAACTTCAATTTCTGGGCCAACTCTCTGGACT 1512

Qy 1278 ATGAAGAAGATTTAAGAACAGACTGGAGCATCAGGGCCAAAGAGCAAACTCAAGCTGC 1337  
Db 1513 ATGAAGAAGATTTAAGAACAGACTGGAGCATCAGGGCCAAAGAGCAAACTCAAGCTGC 1572  
Qy 1338 TGCACCTTGAAGAAGCCAAATGAACCTGTCCCTGTCTGTCTCAGAGGGTGGACAGAAAGCG 1397  
Db 1573 TGCACCTTGAAGAAGCCAAATGAACCTGTCCCTGTCTGTCTCAGAGGGTGGACAGAAAGCG 1632  
Qy 1398 AGAGCCCTCAGTCCACCTGTGTCGCACTCTGCTACCTCTCAGAGGCGAGGACAAAGGC 1457  
Db 1633 AGAGCCCTCAGTCCACCTGTGTCGCACTCTGCTACCTCTCAGAGGCGAGGACAAAGGC 1692  
Qy 1458 CCGTGCAATCCCGCAGCGTGCACCGTGCACCGTGCACCGTGCCTGTAGAGACA 1517  
Db 1693 CCGTGCAATCCCGCAGCGTGCACCGTGCACCGTGCACCGTGCCTGTAGAGACA 1752  
Qy 1518 GCCCGCTGTACAGGCGCTCAGTGGCTGCACCTGTCCGAGACAGGCTGGAGAGACA 1577  
Db 1753 GCCCGCTGTACAGGCGCTCAGTGGCTGCACCTGTCCGAGACAGGCTGGAGAGACA 1812  
Qy 1578 ATAAGCTCAAGCGTTCTCTCTCTGATATCAAAATCAGTTTTCATATTCAGCCAGCATGG 1637  
Db 1813 ATAAGCTCAAGCGTTCTCTCTCTGATATCAAAATCAGTTTTCATATTCAGCCAGCATGG 1872  
Qy 1638 CAGCATCCTTACATGGCTTCTCTCATCAGAGATGCTTTGGAAATCTACAACCTTCCA 1697  
Db 1873 CAGCATCCTTACATGGCTTCTCTCATCAGAGATGCTTTGGAAATCTACAACCTTCCA 1932  
Qy 1698 CTACTCTGATGGGACCAACAAAGCTATGCCAGTCTCTCCCTGTTCAGGAACTATCGAGC 1757  
Db 1933 CTACTCTGATGGGACCAACAAAGCTATGCCAGTCTCTCCCTGTTCAGGAACTATCGAGC 1992  
Qy 1758 AGACTCCGAAACAGTCTCTGATAGGAGAGCCAGCATCCCAAGAGCTGGAGACCG 1817  
Db 1993 AGACTCCGAAACAGTCTCTGATAGGAGAGCCAGCATCCCAAGAGCTGGAGACCG 2052  
Qy 1818 CAGAGCTTTCAGACAGCAGAGAGGATTCGATTCGTCAGAACCCAGCAGAGTGGCA 1877  
Db 2053 CAGAGCTTTCAGACAGCAGAGAGGATTCGATTCGTCAGAACCCAGCAGAGTGGCA 2112  
Qy 1878 CGGCCCAGAGTCCCTTTTATCTCCACTGCAATGAAGTGGAGCGTGGAGACAATTACC 1937  
Db 2113 CGGCCCAGAGTCCCTTTTATCTCCACTGCAATGAAGTGGAGCGTGGAGACAATTACC 2172  
Qy 1938 ACACAGCTTCTTTTCCGGCTTTCCACAGCAGCAGCCTCACAGTCTCTGCGCC 1997  
Db 2173 ACACAGCTTCTTTTCCGGCTTTCCACAGCAGCAGCCTCACAGTCTCTGCGCC 2232  
Qy 1998 TGGGCTTTAAGGGCTGGCACTCGGATATCTTTGGCCCCCAGACCTCTACCCCTTCCCTGA 2057  
Db 2233 TGGGCTTTAAGGGCTGGCACTCGGATATCTTTGGCCCCCAGACCTCTACCCCTTCCCTGA 2292  
Qy 2058 CAGACAGCTGTTATTTGGCAAGAGTCTCACAATCTACTCTGCTCAGCCTACGCTACG 2117  
Db 2293 CCAGCAGCTGTTATTTTGGCAAGAGTCTCACAATCTACTCTGCTCAGCCTACGCTACG 2352  
Qy 2118 GAGGCAAGTGCAGTTCCTCTGCTACAGCTCAGCAGCTGCCACCTTGGAGACCAAG 2177  
Db 2353 GAGGCAAGTGCAGTTCCTCTGCTACAGCTCAGCAGCTGCCACCTTGGAGACCAAG 2412  
Qy 2178 TCTATTCTGTGCGAGGCGGAGAACCAAGTGAACAGTGAACAGTGAACAGTGAACAGTGAAC 2237  
Db 2413 TCTATTCTGTGCGAGGCGGAGAACCAAGTGAACAGTGAACAGTGAACAGTGAACAGTGAAC 2472  
Qy 2238 ATCAGAGAGCCCTTTTGAAGAGAGTGAACAGTGAACAGTGAACAGTGAACAGTGAAC 2297  
Db 2473 ATCAGAGAGCCCTTTTGAAGAGAGTGAACAGTGAACAGTGAACAGTGAACAGTGAAC 2532  
Qy 2298 AGAGCATCATGTGAGAGAACAGGTCACGGGAGAGCTGGGAAAGTGGGAGTGGGAGTCACTCA 2357  
Db 2533 AGAGCATCATGTGAGAGAACAGGTCACGGGAGAGCTGGGAAAGTGGGAGTGGGAGTCACTCA 2592



Db 243 AAAGAATGATGAATGACATGGCTGCTCCAAAGCATCTTTGTTGTGGAAATGGTTATTC 302  
Qy 301 AGTCATCTCTTTATGAATCAAAATGTAGGGGCTGCTTTGTGAGCGAGTCTCTTTGCAAGA 360  
Db 303 AGTCATCTCTTTATGAATCAAAATGTAGGGGCTGCTTTGTGAGCGAGTCTCTTTGCAAGA 362  
Qy 361 GCACATCAACGGGAAAGAGAAAGAGACATTCACCTTGGAGGGCTCTTGCTGAAATGGGTT 420  
Db 363 GCACATCAACGGGAAAGAGAAAGAGACATTCACCTTGGAGGGCTCTTGCTGAAATGGGTT 422  
Qy 421 TAACTCTCTCTTTGCGCAGTCACACAGCGCTGACCTCATACATTTTATAGTCAATGGAGT 480  
Db 423 TAACTCTCTCTTTGCGCAGTCACACAGCGCTGACCTCATACATTTTATAGTCAATGGAGT 482  
Qy 481 GGCTGAGCCTTTGAGCACACCCATTCATCATCTGTCGCAAAATTAAGAAAGAGAGTGGG 540  
Db 483 GGCTGAGCCTTTGAGCACACCCATTCATCATCTGTCGCAAAATTAAGAAAGAGAGTGGG 542  
Qy 541 AAAAGAGGACTTATTTGTCATGCGCCCATGAGATGATTGGAACCTCAAATTTGTTACTGAG 600  
Db 543 AAAAGAGGACTTATTTGTCATGCGCCCATGAGATGATTGGAACCTCAAATTTGTTACTGAG 602  
Qy 601 AGCTTGGTGGCTCTGCTGGAAGTGAACCGAAAGTGTCTTAATCAACTGCTCCAAAGCTT 660  
Db 603 AGCTTGGTGGCTCTGCTGGAAGTGAACCGAAAGTGTCTTAATCAACTGCTCCAAAGCTT 662  
Qy 661 TTTGTGGAATACAAATACATCCACATTTTGGAGGCCATTAATCAACTGCTCCAAAGCTT 720  
Db 663 TTTGTGGAATACAAATACATCCACATTTTGGAGGCCATTAATCAACTGCTCCAAAGCTT 722  
Qy 721 ATGAAGGAAGTGTGAACAGCAAGATGTTAAATACAGAGCTCATCCAGCATTCAGCG 780  
Db 723 ATGAAGGAAGTGTGAACAGCAAGATGTTAAATACAGAGCTCATCCAGCATTCAGCG 782  
Qy 781 AAACATAAGTGTGACATGATTGACGTACAGAGTGTAGTTTACGATCAAGCTCCCAA 840  
Db 783 AAACATAAGTGTGACATGATTGACGTACAGAGTGTAGTTTACGATCAAGCTCCCAA 842  
Qy 841 GATGTTGCTCTCTCTCTTCAGACTGTTTCTCACTGACTTCTGGGTAACACTGAGAG 900  
Db 843 GATGTTGCTCTCTCTCTTCAGACTGTTTCTCACTGACTTCTGGGTAACACTGAGAG 902  
Qy 901 AGCTTCAACTCTGTTTCACTGCTTGCAGTGGGTTGCTGAGTCTCTGTTGTTCCCT 926  
Db 903 AGCTTCAACTCTGTTTCACTGCTTGCAGTGGGTTGCTGAGTCTCTGTTGTTCCCT 962  
Qy 927 ----- 926  
Db 963 GGCTCTGTGAAGAAATCCACTCTAGTCCCTACCTGCAATTTCTCAGCCTTGCTTACCT 1022  
Qy 927 ----- 926  
Db 1023 GTTGCCAAACATTTGGGCCAACCCGAAATCTTCCAAATCTTTATCTTTGGTGGCAGGAGAT 1082  
Qy 927 ----- AGGAGCTGATGACAGAGATGGATGGTATGCTTAAATGCCAGCAAT 976  
Db 1083 GTCTCNAACAGAGCTGATGACAGAGATGGATGGTATGCTTAAATGCCAGCAAT 1142  
Qy 977 ACCTGTCGAAGCCTGACTTATFCCCGAGTCTCATTTCTCGGTGGTGGCTGTGAATGAC 1036  
Db 1143 ACCTGTCGAAGCCTGACTTATFCCCGAGTCTCATTTCTCGGTGGTGGCTGTGAATGAC 1202  
Qy 1037 AGCTTTGTGAGAAATTTTGGCGTGGTGGCAAAATCAGTAGATTTTCATTGAGAAAGCA 1096  
Db 1203 AGCTTTGTGAGAAATTTTGGCGTGGTGGCAAAATCAGTAGATTTTCATTGAGAAAGCA 1262  
Qy 1097 AAAGCCTCCAATGGATGTGTTCTAGTGAATCTGTTAGCTGGGATCTCCCGTCCGCCACC 1156  
Db 1263 AAAGCCTCCAATGGATGTGTTCTAGTGAATCTGTTAGCTGGGATCTCCCGTCCGCCACC 1322  
Qy 1157 ATGCTATCCCTACATCATGAAGAGATGGACATGCTTTTATAGTGAAGCTTACAGATT 1216  
Db 1323 ATGCTATCCCTACATCATGAAGAGATGGACATGCTTTTATAGTGAAGCTTACAGATT 1382

Qy 1217 GTGAAGAAAGAGACCTACTATATCTTCCAACTTCAATTTTCTGGGCCAACTCCTGGAC 1276  
Db 1383 GTGAAGAAAGAGACCTACTATATCTTCCAACTTCAATTTTCTGGGCCAACTCCTGGAC 1442  
Qy 1277 TATGAGAAAGAGATTAAAGAACAGACTGGAGCATCAGGGCCAAAGAGCAAACTCAAGCTG 1336  
Db 1443 TATGAGAAAGAGATTAAAGAACAGACTGGAGCATCAGGGCCAAAGAGCAAACTCAAGCTG 1502  
Qy 1337 CTGACCTGGAGAGCCAAATGAACCTGTCCCTGCTCTCAGAGGGTGGACAGAAAGC 1396  
Db 1503 CTGACCTGGAGAGCCAAATGAACCTGTCCCTGCTCTCAGAGGGTGGACAGAAAGC 1562  
Qy 1397 GAGAGCGCCCTCAGTCCACCTGTGCCACTCTGTACTCTCAGAGGAGCAGAGCAAAAGG 1456  
Db 1563 GAGAGCGCCCTCAGTCCACCTGTGCCACTCTGTACTCTCAGAGGAGCAGAGCAAAAGG 1622  
Qy 1457 CCGGTGCAATCCCGCCAGCGTCCAGCGTCCAGCGTCCAGCGTCCAGCGTCCAGCGTCCAGCG 1516  
Db 1623 CCGGTGCAATCCCGCCAGCGTCCAGCGTCCAGCGTCCAGCGTCCAGCGTCCAGCGTCCAGCG 1682  
Qy 1517 AGCCCGCTGGTACAGCGCTCAGTGGGCTGACCTGTCCGAGACAGGCTGGAGAGACAGC 1576  
Db 1683 AGCCCGCTGGTACAGCGCTCAGTGGGCTGACCTGTCCGAGACAGGCTGGAGAGACAGC 1742  
Qy 1577 AATAAGCTCAAGGTTCTCTCTCTGATATCAAAATCAGTTCATATTCAGCCAGCATG 1636  
Db 1743 AATAAGCTCAAGGTTCTCTCTCTGATATCAAAATCAGTTCATATTCAGCCAGCATG 1802  
Qy 1637 GCAGATCTTACATGGCTCTCTCTCATCAGAGATGCTTTGGAATCTCAAACTTCC 1696  
Db 1803 GCAGATCTTACATGGCTCTCTCTCATCAGAGATGCTTTGGAATCTCAAACTTCC 1862  
Qy 1697 ACTACTCTGGATGGAGCAACAAAGCTATGCGAGTCTCCCTGTTTCAGGAACCTATCGAG 1756  
Db 1863 ACTACTCTGGATGGAGCAACAAAGCTATGCGAGTCTCCCTGTTTCAGGAACCTATCGAG 1922  
Qy 1757 CAGACTCCGAAACAGTCTCTGATGAAGAGAGAGCAGCATCCCCAAGAGCTCGAGAC 1816  
Db 1923 CAGACTCCGAAACAGTCTCTGATGAAGAGAGAGCAGCATCCCCAAGAGCTCGAGAC 1982  
Qy 1817 GCCAGCCTTCAGACGCGCAGAGCAAGCGATTGCAATTCGGTTCAGAACCCAGCAGTGGC 1876  
Db 1983 GCCAGCCTTCAGACGCGCAGAGCAAGCGATTGCAATTCGGTTCAGAACCCAGCAGTGGC 2042  
Qy 1877 ACCGCCCAGAGTCCCTTTTATCTCCTCATCGAAGTGGAGCGTGGAGGACAATTAC 1936  
Db 2043 ACCGCCCAGAGTCCCTTTTATCTCCTCATCGAAGTGGAGCGTGGAGGACAATTAC 2102  
Qy 1937 CACACAGCTTCTTTTCGGCTTTCCACAGCAGCAGCACCTCAGAAAGTCTGTGGC 1996  
Db 2103 CACACAGCTTCTTTTCGGCTTTCCACAGCAGCAGCACCTCAGAAAGTCTGTGGC 2162  
Qy 1997 CTGGCCTTAAAGGCTGCACTCGGATATCTTGCCCGCCAGAGCTCTAGCCCTTCCCTG 2056  
Db 2163 CTGGCCTTAAAGGCTGCACTCGGATATCTTGCCCGCCAGAGCTCTAGCCCTTCCCTG 2222  
Qy 2057 ACCAGAGCTGTATTTTGGCAGAGTCTCAGCTTCTACTCTGCTCAGCCATCTAC 2116  
Db 2223 ACCAGAGCTGTATTTTGGCAGAGTCTCAGCTTCTACTCTGCTCAGCCATCTAC 2282  
Qy 2117 GGAGGAGTCCAGTACTCTGCTTACAGTTCAGCAGCAGCTGCCACTTGGCGAGACCAA 2176  
Db 2283 GGAGGAGTCCAGTACTCTGCTTACAGTTCAGCAGCAGCTGCCACTTGGCGAGACCAA 2342  
Qy 2177 GTCTATTTGTGCGCAGCGGAGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2236  
Db 2343 GTCTATTTGTGCGCAGCGGAGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2402  
Qy 2237 CATGAAGAGAGCCCTTTTGAAGAGCAGTTTAAAGCAGAGCTGCCAAATGGAATTTGGA 2296  
Db 2403 CATGAAGAGAGCCCTTTTGAAGAGCAGTTTAAAGCAGAGCTGCCAAATGGAATTTGGA 2462



Db 601 GGAACGGAAAAAGTGTCTAATTGATAGCCGGCCATTTGTGGAATACAATACATCCAC 660  
Qy 685 ATTTTGGAAAGCCATTAATATCAACTGCTCCAGCTTATGAAGCGAAGTTGCAACAGAC 744  
Db 661 ATTTTGGAAAGCCATTAATATCAACTGCTCCAGCTTATGAAGCGAAGTTGCAACAGAC 720  
Qy 745 AAAGTGTAAATACAGAGTCTATCCAGCATTCAGCGAAACATAGGTTGACATTTGTC 804  
Db 721 AAAGTGTAAATACAGAGTCTATCCAGCATTCAGCGAAACATAGGTTGACATTTGTC 780  
Qy 805 AGTCAGAAAGTTGATTTACGATCAAAAGCTCCCAAGATGTTGCCCTCTCTCTCTCAGAC 864  
Db 781 AGTCAGAAAGTTGATTTACGATCAAAAGCTCCCAAGATGTTGCCCTCTCTCTCTCAGAC 840  
Qy 865 TGTCTTCTCAGCTCTCTCTGTTAACTGGAAGAGAGCTTCAACTCTGTTCACTGCT 924  
Db 841 TGTCTTCTCAGCTCTCTCTGTTAACTGGAAGAGAGCTTCAACTCTGTTCACTGCT 900  
Qy 925 GC----- 926  
Db 901 GCAGGTGGGTTTGTCTGAGTTCTCTCGTTGTTTCCCTGGCCTCTGTGAAGGAAATCCACT 960  
Qy 927 ----- 926  
Db 961 CTAGTCCTACTGCAATTTCTCAGCCTTGTCTTACCTGTTGCCAATGTTGGGCCAACCCGA 1020  
Qy 927 -----AGGAGCTGATGCAG 940  
Db 1021 ATTCTTCCCAATCTTTATCTTGGCTGCCAGCGAGATGCTCTCAACAGAGCTGATGCAG 1080  
Qy 941 CAGAAATGGGATGGTATGTTGTTAAATGCCAGCAATACCTGTCCAAAGCCTGACTTTATC 1000  
Db 1081 CAGAAATGGGATGGTATGTTGTTAAATGCCAGCAATACCTGTCCAAAGCCTGACTTTATC 1140  
Qy 1001 CCCGAGTCTCATTTCTCTGCTGTGCTGTGAATGACAGCTTTTGTGAGAAAATTTTGCG 1060  
Db 1141 CCCGAGTCTCATTTCTCTGCTGTGCTGTGAATGACAGCTTTTGTGAGAAAATTTTGCG 1200  
Qy 1061 TGGTTGACAAATCAGTAGATTTTATGAGAAAGCAAAAGCCTCCAAATGGATGTTCTTA 1120  
Db 1201 TGGTTGACAAATCAGTAGATTTTATGAGAAAGCAAAAGCCTCCAAATGGATGTTCTTA 1260  
Qy 1121 GTGCACCTGTTTACGTGGATCTCCCGCTCCGCCACCATCGCTATCGCTTACATCATGAAG 1180  
Db 1261 GTGCACCTGTTTACGTGGATCTCCCGCTCCGCCACCATCGCTATCGCTTACATCATGAAG 1320  
Qy 1181 AGGATGACATGTCTTTAGATGAAGCTTACAGATTTTGTGAAAAGAAAAGACCTACTATA 1240  
Db 1321 AGGATGACATGTCTTTAGATGAAGCTTACAGATTTTGTGAAAAGAAAAGACCTACTATA 1380  
Qy 1241 TCTCCAAACTCAATTTTCTGGGCCAACTCCTGGACTATGAGAGAGATTAAGAACCCAG 1300  
Db 1381 TCTCCAAACTCAATTTTCTGGGCCAACTCCTGGACTATGAGAGAGATTAAGAACCCAG 1440  
Qy 1301 ACTGGAGCATCAGGGCCAAAGAGCAAACTCAAGCTGCTGCACCTGGAGAGGCCAAATGAA 1360  
Db 1441 ACTGGAGCATCAGGGCCAAAGAGCAAACTCAAGCTGCTGCACCTGGAGAGGCCAAATGAA 1500  
Qy 1361 CCTGTCCCTGTCTCAGAGGGTGGACAGAAAAGGAGAGCGCCCTCAGTCCAACCTGT 1420  
Db 1501 CCTGTCCCTGTCTCAGAGGGTGGACAGAAAAGGAGAGCGCCCTCAGTCCAACCTGT 1560  
Qy 1421 GCCGACTCTGCTACTCTCAGAGGGCAGAGGACAAAGCCCGTGCATCCCGCCAGCGTGCCC 1480  
Db 1561 GCCGACTCTGCTACTCTCAGAGGGCAGAGGACAAAGCCCGTGCATCCCGCCAGCGTGCCC 1620  
Qy 1481 AGCGTGCCAGCGTGCAGCGCTGCTGTTAGAGGACAGCCCGCTGTTAGAGGGCTCAGT 1540  
Db 1621 AGCGTGCCAGCGTGCAGCGCTGCTGTTAGAGGACAGCCCGCTGTTAGAGGGCTCAGT 1680  
Qy 1541 GGGCTGCACCTGTCCGAGAGCGCTGGAGAGCAGCAATTAAGCTCAAGCGTTCTCTCTCT 1600  
Db 1681 GGGCTGCACCTGTCCGAGAGCGCTGGAGAGCAGCAATTAAGCTCAAGCGTTCTCTCTCTCT 1740

Qy 1601 CTGGATATCAAAATCAGTTTCATATTCAGCCAGCATGGCAGCATCCTTTACATGGCTTCTCC 1660  
Db 1741 CTGGATATCAAAATCAGTTTCATATTCAGCCAGCATGGCAGCATCCTTTACATGGCTTCTCC 1800  
Qy 1661 TCATCAGAAAGATGCTTTGGAAATACATAAAACCTTCCACTACTCTGGATGGGACCAACAG 1720  
Db 1801 TCATCAGAAAGATGCTTTGGAAATACATAAAACCTTCCACTACTCTGGATGGGACCAACAG 1860  
Qy 1721 CTATGCCAGTTCTCCCTGTTTCAGAACTATCGAGAGCATCCCGAAACAGAGTCTCTGAT 1780  
Db 1861 CTATGCCAGTTCTCCCTGTTTCAGAACTATCGAGAGCATCCCGAAACAGAGTCTCTGAT 1920  
Qy 1781 AAGGAGGAGCCAGCATCCCAAGAGCTGCAGACCCAGGCTTTCAGACAGCCAGC 1840  
Db 1921 AAGGAGGAGCCAGCATCCCAAGAGCTGCAGACCCAGGCTTTCAGACAGCCAGC 1980  
Qy 1841 AAGCGATTTGCTATTCGGTTCAGAACCCAGAGCAGTGCACCCGCGCAGAGGTCCTTTTATCT 1900  
Db 1981 AAGCGATTTGCTATTCGGTTCAGAACCCAGAGCAGTGCACCCGCGCAGAGGTCCTTTTATCT 2040  
Qy 1901 CCACTGCTATCGAAGTGGGAGCGTGGAGGACAAATTAACCAACAGCTTCTTTTCGGCTT 1960  
Db 2041 CCACTGCTATCGAAGTGGGAGCGTGGAGGACAAATTAACCAACAGCTTCTTTTCGGCTT 2100  
Qy 1961 TCACAGCAGCAGCAGCATCTCAGAACTGCTGGCTGGCTTAAAGGCTGGCAGCTCG 2020  
Db 2101 TCACAGCAGCAGCAGCATCTCAGAACTGCTGGCTGGCTTAAAGGCTGGCAGCTCG 2160  
Qy 2021 GATATCTGGGCCCCCAGACCTCTACCCCTTCCCTGACCCAGCAGCTGTAATTTTGCACA 2080  
Db 2161 GATATCTGGGCCCCCAGACCTCTACCCCTTCCCTGACCCAGCAGCTGTAATTTTGCACA 2220  
Qy 2081 GAGTCTCTACATTTCTCTGCTCTGCTCAGCCATCTACCGAGCAGTGCAGTACTCTGCC 2140  
Db 2221 GAGTCTCTACATTTCTCTGCTCTGCTCAGCCATCTACCGAGCAGTGCAGTACTCTGCC 2280  
Qy 2141 TACAGCTGCAGCAGCAGCTGCCACTTTCGGAGACCAAGTCTATTCTGTGCGCAGCGGCGAG 2200  
Db 2281 TACAGCTGCAGCAGCAGCTGCCACTTTCGGAGACCAAGTCTATTCTGTGCGCAGCGGCGAG 2340  
Qy 2201 AAGCCAAAGTGACAGAGCTGACTTCGCGCGGAGCTGGCATGAAGAGAGCCCTTTGAAAAG 2260  
Db 2341 AAGCCAAAGTGACAGAGCTGACTTCGCGCGGAGCTGGCATGAAGAGAGCCCTTTGAAAAG 2400  
Qy 2261 CAGTTTAAACGAGAGCTGCCAAATGGAATTTGGAGAGCATCATGTCCAGAGAACAGG 2320  
Db 2401 CAGTTTAAACGAGAGCTGCCAAATGGAATTTGGAGAGCATCATGTCCAGAGAACAGG 2460  
Qy 2321 TCACGGGAAAGAGCTGGGAAAGTGGGCGAGTCTAGCTTTTCGGGCGAGCATGGAAATC 2380  
Db 2461 TCACGGGAAAGAGCTGGGAAAGTGGGCGAGTCTAGCTTTTCGGGCGAGCATGGAAATC 2520  
Qy 2381 ATTGAGGTCTCTGAGAGAAAGACATTTGACTTCTATAGACAAATTTTTTTTCTTGT 2440  
Db 2521 ATTGAGGTCTCTGAGAGAAAGACATTTGACTTCTATAGACAAATTTTTTTTCTTGT 2580  
Qy 2441 TCACAAAAAAATTTCCCTGTAAATCTGAAATATATATATATATATATATATATATATTTTGG 2500  
Db 2581 TCACAAAAAAATTTCCCTGTAAATCTGAAATATATATATATATATATATATATATTTTGG 2640  
Qy 2501 AAAATCGAGCTATGTTTAAAGAACAGGTGGATCAACCCAGTTGTTACTCTCTTAACA 2560  
Db 2641 AAAATCGAGCTATGTTTAAAGAACAGGTGGATCAACCCAGTTGTTACTCTCTTAACA 2700  
Qy 2561 TCTGCAATTTGAGAGATCAGCTAATCTCTCTCAAAAAATGGAAGGCGAGATGCTAGA 2620  
Db 2701 TCTGCAATTTGAGAGATCAGCTAATCTCTCTCAAAAAATGGAAGGCGAGATGCTAGA 2760  
Qy 2621 ATCCCCCTTAGACGGAGGAAAAACCAATTTTATTCAGTGAATTAACATCTCTTGTCTTAA 2680  
Db 2761 ATCCCCCTTAGACGGAGGAAAAACCAATTTTATTCAGTGAATTAACATCTCTTGTCTTAA 2820







Db 781 AGTCAGAGGTTGTAGTTTACGATCAAGCTCCCAAGATGTTGCTCTCTCTTTCAGAC 840  
Qy 865 TGTCTTCTCACTGTACTTCTGGGTAACTGAGAGAGAGCTCAACTCTGTTTCACTGCTT 924  
Db 841 TGTCTTCTCACTGTACTTCTGGGTAACTGAGAGAGAGCTCAACTCTGTTTCACTGCTT 900  
Qy 925 GC----- 926  
Db 901 CAGGTGGGTTGCTCAGTTCTCTGTTGTTTCCCTGGCTCTGTGAAGGAAATTCAC 960  
Qy 927 ----- 926  
Db 961 CTAGTCCCTACCTGCAATTTCTCAGCGTTGCTTACCTGTTGCAACATGGGCCAACCGA 1020  
Qy 927 -----AGGAGCTGATCG 940  
Db 1021 ATTCTTCCCAATCTTATCTTGCTGCCAGCAGATGCTCCTCAACAGGAGCTGATACAG 1080  
Qy 941 CAGATGGGATGTTGTTATGTTAAATGCCAGCAATACCTGTGCAAGCCTGACTTTATC 1000  
Db 1081 CAGATGGGATGTTGTTATGTTAAATGCCAGCTATACCTGTGCAAGCCTGACTTTATC 1140  
Qy 1001 CCGAGTCTCAATTTCTGCTGGTGTGCTGTGAATGACAGCTTTGTGAGAAAATTTGCGG 1060  
Db 1141 CCGAGTCTCAATTTCTGCTGGTGTGCTGTGAATGACAGCTTTGTGAGAAAATTTGCGG 1200  
Qy 1061 TGGTTGGACAAATCAGTAGATTTCAATGAGAAAGCAAAAGCCTCCAAATGGATGTTCTTA 1120  
Db 1201 TGGTTGGACAAATCAGTAGATTTCAATGAGAAAGCAAAAGCCTCCAAATGGATGTTCTTA 1260  
Qy 1121 GTGCACTGTTAGCTGGATCTCCGCTCCGCGACATGCTGCTGCTACATCATGAG 1180  
Db 1261 GTGCACTGTTAGCTGGATCTCCGCTCCGCGACATGCTGCTGCTACATCATGAG 1320  
Qy 1181 AGGATGACATGCTTTAGATGAAGCTTACAGATTTGTAAGAAAGAAAGACCTACTATA 1240  
Db 1321 AGGATGACATGCTTTAGATGAAGCTTACAGATTTGTAAGAAAGAAAGACCTACTATA 1380  
Qy 1241 TCTCCAACTTCAATTTCTGGGCCAACTCTGGAATATGAGAGAGAGATTAAGAACAG 1300  
Db 1381 TCTCCAACTTCAATTTCTGGGCCAACTCTGGAATATGAGAGAGAGATTAAGAACAG 1440  
Qy 1301 ACTGGAGCATCAGGCCAAGACCAACTCAGCTGCTGCACTGGAGAGGCCAAATGAA 1360  
Db 1441 ACTGGAGCATCAGGCCAAGACCAACTCAGCTGCTGCACTGGAGAGGCCAAATGAA 1500  
Qy 1361 CTTGTCCCTGCTCTCAGAGGTTGACAGAAAAGCAGACGCCCTCAGTCCACCTGT 1420  
Db 1501 CTTGTCCCTGCTCTCAGAGGTTGACAGAAAAGCAGACGCCCTCAGTCCACCTGT 1560  
Qy 1421 GCCGACTCTGCTACCTCAGAGGACAGAGCAAAAGCCCGTGCATCCGCGCAGGTGCC 1480  
Db 1561 GCCGACTCTGCTACCTCAGAGGACAGAGCAAAAGCCCGTGCATCCGCGCAGGTGCC 1620  
Qy 1481 AGCGTGCCACCGTGCAGCGCTCGCTGTTAGAGGACAGCCCGCTGTAAGCGCTCAGT 1540  
Db 1621 AGCGTGCCACCGTGCAGCGCTCGCTGTTAGAGGACAGCCCGCTGTAAGCGCTCAGT 1680  
Qy 1541 GGGCTGCACTCTGTCGAGACAGAGCTGGAAGACAGCAATTAAGCTCAAGGTTCTTCTCT 1600  
Db 1681 GGGCTGCACTCTGTCGAGACAGAGCTGGAAGACAGCAATTAAGCTCAAGGTTCTTCTCT 1740  
Qy 1601 CTGGATATCAAAATCAGTTTCATATTTCAGCCAGCATGGCAGCATCTTACATGGCTTCTCC 1660  
Db 1741 CTGGATATCAAAATCAGTTTCATATTTCAGCCAGCATGGCAGCATCTTACATGGCTTCTCC 1800  
Qy 1661 TCATCAGAGATGCTTTGGAACTACTACAACTTCCACTACTCTGGATGGGACCAACAG 1720  
Db 1801 TCATCAGAGATGCTTTGGAACTACTACAACTTCCACTACTCTGGATGGGACCAACAG 1860  
Qy 1721 CTATGCCAGTCTTCCCTGTTTCAGGAACATATCGGACGACATCCCGAACACGCTCTGAT 1780  
Db 1861 CTATGCCAGTCTTCCCTGTTTCAGGAACATATCGGACGACATCCCGAACACGCTCTGAT 1920

Qy 1781 AAGGAGGAGCCAGCATCCCAAGAGCTGAGACCCGACGGCCTTCAGACAGCCAGAC 1840  
Db 1921 AAGGAGGAGCCAGCATCCCAAGAGCTGAGACCCGACGGCCTTCAGACAGCCAGAC 1980  
Qy 1841 AAGCCATTTGCATTTGGTTCAGAACCCAGCAGCAGTGGCACCCGCCAGAGGTCCTTTTATCT 1900  
Db 1981 AAGCCATTTGCATTTGGTTCAGAACCCAGCAGCAGTGGCACCCGCCAGAGGTCCTTTTATCT 2040  
Qy 1901 CCATGTCATCGAAGTGGAGCGTGGAGACAAATTAACACACCAGCTTCCTTTTCGGGCTT 1960  
Db 2041 CCATGTCATCGAAGTGGAGCGTGGAGACAAATTAACACACCAGCTTCCTTTTCGGGCTT 2100  
Qy 1961 TCCACAGCAGCAGCAGCACTCAAGAGCTCTGCTGGCCTTAAGGGCTGAGGGCTGSCACTCG 2020  
Db 2101 TCCACAGCAGCAGCAGCACTCAAGAGCTCTGCTGGCCTTAAGGGCTGAGGGCTGSCACTCG 2160  
Qy 2021 GATATCTTTGGCCCCCAGACCTCTACCCCTTCCCTGACAGCAGCTGTTATTTGCCACA 2080  
Db 2161 GATATCTTTGGCCCCCAGACCTCTACCCCTTCCCTGACAGCAGCTGTTATTTGCCACA 2220  
Qy 2081 GAGTCTCTCACACTTCTACTCTGCTCAGCCTATACGAGGAGCAGTGCAGTTACTCTGCC 2140  
Db 2221 GAGTCTCTCACACTTCTACTCTGCTCAGCCTATACGAGGAGCAGTGCAGTTACTCTGCC 2280  
Qy 2141 TACAGCTGAGCAGCTGCCACTTGGGAGACCAAGTCTAATCTGTGCGCAGGCGCAG 2200  
Db 2281 TACAGCTGAGCAGCTGCCACTTGGGAGACCAAGTCTAATCTGTGCGCAGGCGCAG 2340  
Qy 2201 AAGCCAAGTGCAGAGCTGACTCGCGCGGAGCTGGCATGAAGAGAGCCCTTTGAAAAG 2260  
Db 2341 AAGCCAAGTGCAGAGCTGACTCGCGCGGAGCTGGCATGAAGAGAGCCCTTTGAAAAG 2400  
Qy 2261 CAGTTTAAACGAGAGCTGCGAAATGGAATTTGGAGAGCAGTCAATGTGAGAGAACAG 2320  
Db 2401 CAGTTTAAACGAGAGCTGCGAAATGGAATTTGGAGAGCAGTCAATGTGAGAGAACAG 2460  
Qy 2321 TACCGGAGAGCTGGGAAAGTGGGAGTCAAGTCTAGCTTTTCGGCAGCAGTGGAAATC 2380  
Db 2461 TACCGGAGAGCTGGGAAAGTGGGAGTCAAGTCTAGCTTTTCGGCAGCAGTGGAAATC 2520  
Qy 2381 ATTGAGGCTCTCTGAGAGAAAGACACTTGTGAGCTTCTATAGACAA-TTTTTTTTCTTG 2439  
Db 2521 ATTGAGGCTCTCTGAGAGAAAGACACTTGTGAGCTTCTATAGACAA-TTTTTTTTCTTG 2580  
Qy 2440 TTCACAAAAAATTCCTGTAAATCTGAAATATATATATATATATATATATATATTTTG 2499  
Db 2581 TTCACAAAAAATTCCTGTAAATCTGAAATATATATATATATATATATATATATTTTG 2640  
Qy 2500 GAAAATGGAGCTATGCTGTAAAGCAACAGGTGATCAACCCAGTGTGTTACTCTCTTAAC 2559  
Db 2641 GAAAATGGAGCTATGCTGTAAAGCAACAGGTGATCAACCCAGTGTGTTACTCTCTTAAC 2700  
Qy 2560 ATCTGCATTTGAGAGATCAGCTTAATCTTCTCTCAAAAAATGGAGGGGAGATGCTAG 2619  
Db 2701 ATCTGCATTTGAGAGATCAGCTTAATCTTCTCTCAAAAAATGGAGGGGAGATGCTAG 2760  
Qy 2620 AATCCCTCTAGACGGAGAAAACCAATTTTATTCAGTGAATTAACATCTCTCTTGTCTT 2679  
Db 2761 AATCCCTCTAGACGGAGAAAACCAATTTTATTCAGTGAATTAACATCTCTCTTGTCTT 2820  
Qy 2680 AAAAAAGCAAGTGTCTTTGGTGTGGAGGACAAAATCCCTTACCATTTTTCCACGTTGTC 2739  
Db 2821 AAAAAAGCAAGTGTCTTTGGTGTGGAGGACAAAATCCCTTACCATTTTT-CACGTTGTC 2879  
Qy 2740 TACTAAGAGATCTCAAAATATAGTCTTTGTCGGACCTTCCATAGTACACCTTAGCGCT 2799  
Db 2880 TACTAAGAGATCTCAAAATATAGTCTTTGTCGGACCTTCCATAGTACACCTTAGCGCT 2939  
Qy 2800 GAGACTGAGCCAGCTTGGGGGTGAGGTAGTACCCCTGTTAGGAGACAGAGCTTAGTGGT 2859  
Db 2940 GAGACTGAGCCAGCTTGGGGGTGAGGTAGTACCCCTGTTAGGAGACAGAGCTTAGTGGT 2999



Db 961 CTAGTCCTACCTGCAATTTCTCAGCCTTGCTTACCTGTGTCACATTTGGCCAAACCGA 1020  
Qy 927 -----AGGAGCTGATGAG 940  
Db 1021 ATTCTTCCCAATCTTTATCTTGGCTGCCAGCAGATGCTCTCAACAAGAGCTGATACAG 1080  
Qy 941 CAGAAATGGGATTTGGTTATGTTTAATGCCAGCAATACCTGTCCAAAGCCCTGACTTTATC 1000  
Db 1081 CAGAAATGGGATTTGGTTATGTTTAATGCCAGCTATACCTGTCCAAAGCCCTGACTTTATC 1140  
Qy 1001 CCGAGTCTCATTTCTCGGTGTCCTGGAATGACAGCTTTTGTGAGAAAATTTGCGG 1060  
Db 1141 CCGAGTCTCATTTCTCGGTGTCCTGGAATGACAGCTTTTGTGAGAAAATTTGCGG 1200  
Qy 1061 TGGTTGGACAAATCAGTAGATTTTCAATTTGGAAGCAGAAAGCCCTCAATGGATGTTCTTA 1120  
Db 1201 TGGTTGGACAAATCAGTAGATTTTCAATTTGGAAGCAGAAAGCCCTCAATGGATGTTCTTA 1260  
Qy 1121 GTGCACCTGTTAGCTGGGATCTCCGCTCCGCCACCATCGCTATCGCTACATCATGAAG 1180  
Db 1261 GTGCACCTGTTAGCTGGGATCTCCGCTCCGCCACCATCGCTATCGCTACATCATGAAG 1320  
Qy 1181 AGAATGGACATGCTTTTATGATGAGCTTACAGATTTGTGAAGAAAAGACCTACTATA 1240  
Db 1321 AGAATGGACATGCTTTTATGATGAGCTTACAGATTTGTGAAGAAAAGACCTACTATA 1380  
Qy 1241 TCTCCAAACTTCAATTTCTGGGCCAACTCTCGACTATGAGAAGAAATTAAGAACCAG 1300  
Db 1381 TCTCCAAACTTCAATTTCTGGGCCAACTCTCGACTATGAGAAGAAATTAAGAACCAG 1440  
Qy 1301 ACTGGAGCATCAGGGCCAAAGAGCAAACTCAAGCTGTGCACTGGAGAGCCAAATGAA 1360  
Db 1441 ACTGGAGCATCAGGGCCAAAGAGCAAACTCAAGCTGTGCACTGGAGAGCCAAATGAA 1500  
Qy 1361 CTTGTCCCTCTCTCTCAGAGGGTGACAGAAAAGCAGAGCCCTCAGTCCACCTGT 1420  
Db 1501 CTTGTCCCTCTCTCTCAGAGGGTGACAGAAAAGCAGAGCCCTCAGTCCACCTGT 1560  
Qy 1421 GCCGACTCTGCTACCTCAGAGGCAGCAGCAAAAGCCCGTGCATCCGCCCAGGTGCC 1480  
Db 1561 GCCGACTCTGCTACCTCAGAGGCAGCAGCAAAAGCCCGTGCATCCGCCCAGGTGCC 1620  
Qy 1481 AGGTGCCAGCTGACGCGCTGCTGTGTAGAGGACAGCCCGCTGGTACAGCGCTCAGT 1540  
Db 1621 AGGTGCCAGCTGACGCGCTGCTGTGTAGAGGACAGCCCGCTGGTACAGCGCTCAGT 1680  
Qy 1541 GGGCTGACCTGTCCGACAGACAGCTGGAAGACAGCAATAAGCTCAAGCGTTTCTTCTCT 1600  
Db 1681 GGGCTGACCTGTCCGACAGACAGCTGGAAGACAGCAATAAGCTCAAGCGTTTCTTCTCT 1740  
Qy 1601 CTGGATATCAAAATCAGTTTCATATTAGCCAGATGGCAGCATCCTTACATGGCTTCTCC 1660  
Db 1741 CTGGATATCAAAATCAGTTTCATATTAGCCAGCATGGCAGCATCCTTACATGGCTTCTCC 1800  
Qy 1661 TCATCAGAAGATGCTTTGGAATACTCAAACTTCCACTACTCTGGATGGGACCAACAAG 1720  
Db 1801 TCATCAGAAGATGCTTTGGAATACTCAAACTTCCACTACTCTGGATGGGACCAACAAG 1860  
Qy 1721 CTATGCCAGTTCTCCCTGTTTCAGGAACCTATCGGACAGACTCCCGAAACCAAGTCTGAT 1780  
Db 1861 CTATGCCAGTTCTCCCTGTTTCAGGAACCTATCGGACAGACTCCCGAAACCAAGTCTGAT 1920  
Qy 1781 AAGGAGNAGCCAGCATCCCAAGAGCTGCAGACCCGAGCGCTTCAGACAGCCAGAGC 1840  
Db 1921 AAGGAGNAGCCAGCATCCCAAGAGCTGCAGACCCGAGCGCTTCAGACAGCCAGAGC 1980  
Qy 1841 AAGCGATTGCAATTCGGTCAGAACCCAGCAGCAGTGGCACCCGCCCAGAGGTCCCTTTATCT 1900  
Db 1981 AAGCGATTGCAATTCGGTCAGAACCCAGCAGCAGTGGCACCCGCCCAGAGGTCCCTTTATCT 2040  
Qy 1901 CCATGTCATCGAAGTGGAGGTGGAGGACAAATACCAACCAAGCTTCCTTTTTCGCGCTT 1960  
Db 2041 CCATGTCATCGAAGTGGAGGTGGAGGACAAATACCAACCAAGCTTCCTTTTTCGCGCTT 2100

Qy 1961 TCACACGACGACGACACCTCAGAACTCTGCTGGCCTTGGCCTTAAAGGCTGSCACTCG 2020  
Db 2101 TCACACGACGACGACACCTCAGAACTCTGCTGGCCTTGGCCTTAAAGGCTGSCACTCG 2160  
Qy 2021 GATATCTTTGGGCCCCCAGACACCTCTACCCCTTCCCTGTACCCAGCAGCTGGTATTTTGCACA 2080  
Db 2161 GATATCTTTGGGCCCCCAGACACCTCTACCCCTTCCCTGTACCCAGCAGCTGGTATTTTGCACA 2220  
Qy 2081 GAGTCTCTACACTTCTACTCTGCTCAGCCTATCTACCGAGCAGTGCAGTACTCTGCC 2140  
Db 2221 GAGTCTCTACACTTCTACTCTGCTCAGCCTATCTACCGAGCAGTGCAGTACTCTGCC 2280  
Qy 2141 TACAGCTGACGACGCTGCCACTTGGGAGACCAAGTCTATTCTGTGCGCAGCGCGCAG 2200  
Db 2281 TACAGCTGACGACGCTGCCACTTGGGAGACCAAGTCTATTCTGTGCGCAGCGCGCAG 2340  
Qy 2201 AAGCAGAGTGAACAGAGCTGACTGCGGCGGAGCTGGCATGAAGAGAGCCCTTTGAAAAG 2260  
Db 2341 AAGCAGAGTGAACAGAGCTGACTGCGGCGGAGCTGGCATGAAGAGAGCCCTTTGAAAAG 2400  
Qy 2261 CAGTTTAAACGACAGAGCTGCCAAATGGAATTTGGAGAGCATCATGTCAAGAAACAGG 2320  
Db 2401 CAGTTTAAACGACAGAGCTGCCAAATGGAATTTGGAGAGCATCATGTCAAGAAACAGG 2460  
Qy 2321 TCACGCGAAGAGCTCGGAAAAGTGGCAGTCTAGCTTTTCGGGACGATGGAATC 2380  
Db 2461 TCACGCGAAGAGCTCGGAAAAGTGGCAGTCTAGCTTTTCGGGACGATGGAATC 2520  
Qy 2381 ATTGAGTCTCTCTGAGAGAAAGACACTTGTGACTCTATAGACAA-TTTTTTTTCTTG 2439  
Db 2521 ATTGAGTCTCTCTGAGAGAAAGACACTTGTGACTCTATAGACAA-TTTTTTTTCTTG 2580  
Qy 2440 TTCACAAAAAATTTCCCTGTAAATCTGAAATATATATATGTACATATATATTTTG 2499  
Db 2581 TTCACAAAAAATTTCCCTGTAAATCTGAAATATATATATGTACATATATATTTTG 2640  
Qy 2500 GAAATGGAGCTATGGTGTAAAGCAACAGGTGATCAACCCAGTTGTACTCTCTTAAC 2559  
Db 2641 GAAATGGAGCTATGGTGTAAAGCAACAGGTGATCAACCCAGTTGTACTCTCTTAAC 2700  
Qy 2560 ATCTGCAATTTGAGAGATCAGCTAATCTCTCTCAACAAAAATGGAAGGCGAGATGCTAG 2619  
Db 2701 ATCTGCAATTTGAGAGATCAGCTAATCTCTCTCAACAAAAATGGAAGGCGAGATGCTAG 2760  
Qy 2620 AATCCCTCTAGACGAGGAAACCAATTTTATCTAGTGAATTACATCTCTTGTCTT 2679  
Db 2761 AATCCCTCTAGACGAGGAAACCAATTTTATCTAGTGAATTACATCTCTTGTCTT 2820  
Qy 2680 AAAAAAGCAAGTGTCTTTGCTGTGGAGCAAAATCCCTTACCATTTCACCGTTGTGC 2739  
Db 2821 AAAAAAGCAAGTGTCTTTGCTGTGGAGCAAAATCCCTTACCATTTCACCGTTGTGC 2879  
Qy 2740 TACTAAGAGATCTCAAAATATAGTCTTTGTCCGACCCCTTCCATAGTACACTTAGCGCT 2799  
Db 2880 TACTAAGAGATCTCAAAATATAGTCTTTGTCCGACCCCTTCCATAGTACACTTAGCGCT 2939  
Qy 2800 GAGACTGACCGAGCTTTGGGGTCAAGTAGGTAGACCTGTTAGGAGACAGAGCTAGTGGT 2859  
Db 2940 GAGACTGACCGAGCTTTGGGGTCAAGTAGGTAGACCTGTTAGGAGACAGAGCTAGTGGT 2999  
Qy 2860 AAATCCAGAGAAATGATCTCTATCCAAAGCTGATTCACAAAACCCACGCTACCTGACAGC 2919  
Db 3000 AAATCCAGAGAAATGATCTCTATCCAAAGCTGATTCACAAAACCCACGCTACCTGACAGC 3059  
Qy 2920 CGAGGAGACGAGCATCACTCTGCTGGAGCGNCAATTAGGGGCTTGGCAGGCTTACCT 2979  
Db 3060 CGAGGAGACGAGCATCACTCTGCTGGAGCGNCAATTAGGGGCTTGGCAGGCTTACCT 3119  
Qy 2980 TAGAGCAAAACCCAGTACCTCAGACAGGAAAGTGGGGCTTTGACCACTACCATATCTGGT 3039  
Db 3120 TAGAGCAAAACCCAGTACCTCAGACAGGAAAGTGGGGCTTTGACCACTACCATATCTGGT 3179



1233 CTACTATATCTCAAACTTCAATTTTCTGGGCCAACTCCTGACTATGAGAGAGATTG 1292  
1424 CTACTATATCTCAAACTTCAATTTTCTGGGCCAACTCCTGACTATGAGAGAGATTG 1483  
1293 AGAACAGACTGAGCATCAGGCGCAAGAGCAAACTCAAGCTGCTGACCTGGAGAGC 1352  
1484 AGAACAGACTGAGCATCAGGCGCAAGAGCAAACTCAAGCTGCTGACCTGGAGAGC 1543  
1353 CAAATGAACCTGTCCCTGTCTCTCAGAGGGTGGACAGAAAAGCAGAGCGCCCTCAGTC 1412  
1544 CAAATGAACCTGTCCCTGTCTCTCAGAGGGTGGACAGAAAAGCAGAGCGCCCTCAGTC 1603  
1413 CACCTGTGCCGATCTGTCTACCTCAGAGGACAGGAGCAAAAGGCGCTGCATCCGCGCA 1472  
1604 CACCTGTGCCGATCTGTCTACCTCAGAGGACAGGAGCAAAAGGCGCTGCATCCGCGCA 1663  
1473 GCGTGCCAGCGTGCCAGCGTGCAGCGCTCGCTGTGTAGAGACAGCCCGCTGGTACAGG 1532  
1664 GCGTGCCAGCGTGCCAGCGTGCAGCGCTCGCTGTGTAGAGACAGCCCGCTGGTACAGG 1723  
1533 CGCTCAGTGGCTGCACCTCTCCGCGACAGAGCTGGAAGACAGCAATAGCTCAAGCGTT 1592  
1724 CGCTCAGTGGCTGCACCTCTCCGCGACAGAGCTGGAAGACAGCAATAGCTCAAGCGTT 1783  
1593 CCTTCTCTCTGGATATCAAAATCAGTTTCTATTTACGCGAGATGCGAGCATCCTTACATG 1652  
1784 CCTTCTCTCTGGATATCAAAATCAGTTTCTATTTACGCGAGATGCGAGCATCCTTACATG 1843  
1653 GCTTCTCTCTCAGAGATGCTTTGGAACTACTACAAACCTTCCACTACTCTGGATGGGA 1712  
1844 GCTTCTCTCTCAGAGATGCTTTGGAACTACTACAAACCTTCCACTACTCTGGATGGGA 1903  
1713 CCAACAGCTATCCGATCTCCCTCTGTCAGAACTATCGGACAGATCCCGGAAACCA 1772  
1904 CCAACAGCTATCCGATCTCCCTCTGTCAGAACTATCGGACAGATCCCGGAAACCA 1963  
1773 GTCTGTAAAGAGAGAGCCAGCATCCCAAGAGCTGCAGACCGCCAGCGCTTCCAGACA 1832  
1964 GTCTGTAAAGAGAGAGCCAGCATCCCAAGAGCTGCAGACCGCCAGCGCTTCCAGACA 2023  
1833 GCCAGACAGAGATTCGATTCGGTCAGAACAGAGCAGTGGCAGCCGCGAGAGTCCC 1892  
2024 GCCAGACAGAGATTCGATTCGGTCAGAACAGAGCAGTGGCAGCCGCGAGAGTCCC 2083  
1893 TTTTATCTCCACTGCACTGGAAGTGGAGCGTGGAGCAAAATACACACCAAGCTTCCCTTT 1952  
2084 TTTTATCTCCACTGCACTGGAAGTGGAGCGTGGAGCAAAATACACACCAAGCTTCCCTTT 2143  
1953 TCGGCTTTTCCAGCAGCAGAGCACTCAGAAAGTCTGCTGGCTGGCCTTTAAGGGCT 2012  
2144 TCGGCTTTTCCAGCAGCAGAGCACTCAGAAAGTCTGCTGGCTGGCCTTTAAGGGCT 2203  
2013 GGCACTCGGATATCTTGGCCCCCAGACCTTACCCCTTCCCTGACAGCAGCTGGTATT 2072  
2204 GGCACTCGGATATCTTGGCCCCCAGACCTTACCCCTTCCCTGACAGCAGCTGGTATT 2263  
2073 TTGCCACAGAGTCTCTACACTTTCTACTGCTCAGCCATCTAAGAGCAGTGCAGTT 2132  
2264 TTGCCACAGAGTCTCTACACTTTCTACTGCTCAGCCATCTAAGAGCAGTGCAGTT 2323  
2133 ACTTGTCTACAGCTGAGCGAGCTGCCACTTTGGGAGACCAAGTCTATTCTGTGGCGCA 2192  
2324 ACTTGTCTACAGCTGAGCGAGCTGCCACTTTGGGAGACCAAGTCTATTCTGTGGCGCA 2383  
2193 GCGGCGAGAGCCCAAGTGCAGAGCTGACTTCGGGGGAGCTGGCATGAAGAGAGCCCTT 2252  
2384 GCGGCGAGAGCCCAAGTGCAGAGCTGACTTCGGGGGAGCTGGCATGAAGAGAGCCCTT 2443  
2253 TTGAAAAGCAGTTTAAACGAGAGAGCTGCCAAATGGAAATTTGGAGAGAGCATCTGTGAG 2312  
2444 TTGAAAAGCAGTTTAAACGAGAGAGCTGCCAAATGGAAATTTGGAGAGAGCATCTGTGAG 2503

2313 AGAACAGCTCAGCGGAGAGAGCTGGGAAAGTGGGAGTCTAGCTTTTCGGGAGCA 2372  
2504 AGAACAGCTCAGCGGAGAGAGCTGGGAAAGTGGGAGTCTAGCTTTTCGGGAGCA 2563  
2373 TGGAAATCAATTCAGGCTCTCCCTGAGAGAAAGACACTTGTGACTTCTATAGACAATTTTT 2432  
2564 TGGAAATCAATTCAGGCTCTCCCTGAGAGAAAGACACTTGTGACTTCTATAGACAATTTTT 2623  
2433 TTTCTTTGTTCAAAAAAATTTCCCTGTAATCTGAAATATATATATATATATATATATAT 2492  
2624 TTTCTTTGTTCAAAAAAATTTCCCTGTAATCTGAAATATATATATATATATATATATAT 2693  
2493 ATTTTGGAAAAATGAGCTATGCTGTAAGAACCAAGGTGGATCAAGCCAGTTTGTACTC 2552  
2684 ATTTTGGAAAAATGAGCTATGCTGTAAGAACCAAGGTGGATCAAGCCAGTTTGTACTC 2743  
2553 TCTTAACTCTCATTTGAGAGATCAGCTATATATCTTCTCAACAAAAATGGAAGGGCAG 2612  
2744 TCTTAACTCTCATTTGAGAGATCAGCTATATCTTCTCAACAAAAATGGAAGGGCAG 2803  
2613 ATGCTAGATCCCTCCTAGACGAGGAGAAACCATTTTATTCAGTGAATTACACATCTCT 2672  
2804 ATGCTAGATCCCTCCTAGACGAGGAGAAACCATTTTATTCAGTGAATTACACATCTCT 2863  
2673 TGTCTTAAAAAAGCAGTGTCTTTGGTGTGGAGGACAAAAATCCCTTACCATTTCCAC 2732  
2864 TGTCTTAAAAAAGCAGTGTCTTTGGTGTGGAGGACAAAAATCCCTTACCATTTT-CAC 2922  
2733 GTTGTGCTACTAAGAGATCTCAAAATATTAGTCTTTGTCGGGACCTTCCATAGTACACCT 2792  
2923 GTTGTGCTACTAAGAGATCTCAAAATATTAGTCTTTGTCGGGACCTTCCATAGTACACCT 2982  
2793 TAGCGCTGAGACTGAGCCAGCTTTGGGGGTGAGGTAGGTAGGTAGGTAGGTAGGTAGGTAG 2852  
2983 TAGCGCTGAGACTGAGCCAGCTTTGGGGGTGAGGTAGGTAGGTAGGTAGGTAGGTAGGTAG 3042  
2853 TAGTGTAAATCCAGAGAAATGATCTTATCCAAAGCTGATTCACAAACCCACACCTCACC 2912  
3043 TAGTGTAAATCCAGAGAAATGATCTTATCCAAAGCTGATTCACAAACCCACACCTCACC 3102  
2913 TGACAGCCGAGGACACGAGCATCACTCTGTGTGAGCGGACCATTTAGGGGCTTGCCTAAGG 2972  
3103 TGACAGCCGAGGACACGAGCATCACTCTGTGTGAGCGGACCATTTAGGGGCTTGCCTAAGG 3162  
2973 TCTACCTTAGAGAAACCCAGTACCTCAGACAGAAAGTCCGGGCTTTGACCACTACCAT 3032  
3163 TCTACCTTAGAGAAACCCAGTACCTCAGACAGAAAGTCCGGGCTTTGACCACTACCAT 3222  
3033 ATCTGTAGCCCATTTTCTAGGCAATTTGTAATAGGTAGGTAGGTAGGTAGGTAGGTAGGTAG 3092  
3223 ATCTGTAGCCCATTTTCTAGGCAATTTGTAATAGGTAGGTAGGTAGGTAGGTAGGTAGGTAG 3282  
3093 CCAATTCAAACTGTCTATGCAAAAAATTTCCCGTGGGCTTAGATGGAGATAATTTTTTTTT 3152  
3283 CCAATTCAAACTGTCTATGCAAAAAATTTCCCGTGGGCTTAGATGGAGATAATTTTTTTTT 3342  
3153 CTTCTCAGCTTTATGAGAGAGGAGAACTGCTAGGATTCAGCTGGAACCCACAGGACCC 3212  
3343 CTTCTCAGCTTTATGAGAGAGGAGAACTGCTAGGATTCAGCTGGAACCCACAGGACCC 3402  
3213 TGGCAACATCAGGATTTAAGCTAAGTTGGGAGGCTAAGCGAGTCTACCTCCCTCTTTGTA 3272  
3403 TGGCAACATCAGGATTTAAGCTAAGTTGGGAGGCTAAGCGAGTCTACCTCCCTCTTTGTA 3462  
3273 AATCAAAGAAATTTTAAATGGGATTTGCAATCTTTTAAATAAAGATGAACCTTGGTTTC 3332  
3463 AATCAAAGAAATTTTAAATGGGATTTGCAATCTTTTAAATAAAGATGAACCTTGGTTTC 3522

RESULT 9  
AX482372  
LOCUS  
DEFINITION

AX482372 5111 bp DNA linear  
Sequence 41 from Patent WO02057460.

PAT 16-AUG-2002



1905	Db		GCACAGCAACGCGAATGCAATTCCGTGAGAACACAGCAGTGGCGACCCGCCAGAGGTCCTC	1964
1893	QY	TTTTATCTCCACTCGACTCGAAGTGGAGCGGTGGAGACAAATTACCAACACAGCTTCCTTTT		1952
1965	Db	TTTTATCTCCACTCGACTCGAAGTGGAGCGGTGGAGACAAATTACCAACACAGCTTCCTTTT		2024
1953	QY	TCGGCCCTTTCACACAGCCAGCAGCACTCAGCAAGTCTCTGGCCTGGCCCTTAAGGCT	2012	
2025	Db	TCGGCCCTTTCACACAGCCAGCAGCACTCAGCAAGTCTCTGGCCTGGCCCTTAAGGCT	2084	
2013	QY	GGCACTCGGATATCTTGGCCCCCAGACCTCTACCCCTTCCCTGACACAGCTGCTATT	2072	
2085	Db	GGCACTCGGATATCTTGGCCCCCAGACCTCTACCCCTTCCCTGACACAGCTGCTATT	2144	
2073	QY	TTGCGACAGAGTCTCACAATTCTCTCTGCTCAGCCATCTACGGAGCAGTGCAGATT	2132	
2145	Db	TTGCGACAGAGTCTCACAATTCTCTCTGCTCAGCCATCTACGGAGCAGTGCAGATT	2204	
2133	QY	ACTCTGCCCTACAGCTGCAGCCAGCTGCCCACTTGGGAGACCAAGTCTATTCTGTGGCCA	2192	
2205	Db	ACTCTGCCCTACAGCTGCAGCCAGCTGCCCACTTGGGAGACCAAGTCTATTCTGTGGCCA	2264	
2193	QY	GGCGCGAGAAGCCAAAGTGCAGAGCTGACTCGCGGGAGAGCTGGCATCAAGAGAGCCCT	2252	
2265	Db	GGCGCGAGAAGCCAAAGTGCAGAGCTGACTCGCGGGGGAGCTGGCATCAAGAGAGCCCT	2324	
2253	QY	TTGAAAAGCAGTTTAAACGAGAAAGCTGCCAAATGGAAATTGGAGAGAGCATCATGTGAG	2312	
2325	Db	TTGAAAAGCAGTTTAAACGAGAAAGCTGCCAAATGGAAATTGGAGAGAGCATCATGTGAG	2384	
2313	QY	AGAACAGGTCAACGGGAAGAGCTGGGAAAGTGGGCAGTCAAGTCTAGCTTTTCGGGACGCA	2372	
2385	Db	AGAACAGGTCAACGGGAAGAGCTGGGAAAGTGGGCAGTCAAGTCTAGCTTTTCGGGACGCA	2444	
2373	QY	TGGAAAATCATTTGAGGCTCTCCTGAGAGAAAGACACTGTGAGCTTCTATAGACAATTTTT	2432	
2445	Db	TGGAAAATCATTTGAGGCTCTCCTGAGAGAAAGACACTGTGAGCTTCTATAGACAATTTTT	2504	
2433	QY	TTTCTTGTTTCAAAAAAATTCCTGTAAATCTGAAATATATATATGATACATATATAT	2492	
2505	Db	TTTCTTGTTTCAAAAAAATTCCTGTGGAATCTGAAATATATATGATGAGCATATATAT	2564	
2493	QY	ATTTTGTGAAAATGGAGCTATGGTGTAAAGCAACAGGTGGATCAACCCAGTTGTTACTC	2552	
2565	Db	ATTTTGTGAAAATGGAGCTATGGTGTAAAGCAACAGGTGGATCAACCCAGTTGTTACTC	2624	
2553	QY	TCTTAAACATCTGCATTTTGAAGATCAGCTAAATCTTCTCTCAACAAAATGGAGGGCA	2611	
2625	Db	TCTTAAACATCTGCATTTTGAAGATCAGCTAAATCTTCTCTCAACAAAATGGAGGGCA	2684	
2612	QY	GATGCTAGAAATCCCCCTTAGACGAGGAGAAAACCATTTTATTCAGTGAATTAACAATCCTC	2671	
2685	Db	GATGCTAGAAATCCCCCTTAGACGAGGAGAAAACCATTTTATTCAGTGAATTAACAATCCTC	2744	
2672	QY	TGTGTTCTTAAAAAAGCAAGTGTCTTTGGTGTGGAGGACAAAATCCCTACCATTTTCCA	2731	
2745	Db	TGTGTTCTTAAAAAAGCAAGTGTCTTTGGTGTGGAGGACAAAATCCCTACCATTTTCA	2803	
2732	QY	CGTGTGCTACTAAGAGATCTCAAAATATTAGTCTTTGTCCGGAGCCCTTCATAGTACAC	2791	
2804	Db	CGTGTGCTACTAAGAGATCTCAAAATATTAGTCTTTGTCCGGAGCCCTTCATAGTACAC	2863	
2792	QY	TTAGCGCTGAGACTGAGCCAGCTTGGGGGTGAGGTAGGAGCCCTGTTAGGGACAGAGC	2851	
2864	Db	TTAGCGCTGAGACTGAGCCAGCTTGGGGGTGAGGTAGGAGCCCTGTTAGGGACAGAGC	2923	
2852	QY	CTAGTGTGTAATCCAGAGAAATGATCTTATCCAAAGCTGATTACAAACCCACGCTCAC	2911	
2924	Db	CTAGTGTGTAATCCAGAGAAATGATCTTATCCAAAGCTGATTACAAACCCACGCTCAC	2983	
2912	QY	CTGACAGCCGAGGACACGAGCATCACTCTGCTGGACGAGCCATTAGGGGCTTGTGCAAG	2971	

Db	2984	CTGACAGCCGAGGACACGAGCATCACTCTGCTGGACGGACCAATTAGGGGGCTTGCCAG	3043
QY	2972	GTCTACCTTAGAGCAACCCAGTACCTCAGACAGGAAAGTCGGGGCTTTTGACCACTACCA	3031
Db	3044	GTCTACCTTAGAGCACAACCCAGTACCTCAGACAGGAAAGTCGGGGCTTTTGACCACTACCA	3103
QY	3032	TATCTGGTAGCCATTTCTTAGGCATTGTGAATAGTAGTAGGTAGCTAGTCACACATTTTCAG	3091
Db	3104	TATCTGGTAGCCATTTCTTAGGCATTGTGAATAGTAGTAGGTAGCTAGTCACACATTTTCAG	3163
QY	3092	ACCAATTCAAACTGTCTATGTCACAAAATTCCTCGGGCTAGATGGAGATAAATTTTTTT	3151
Db	3164	ACCAATTCAAACTGTCTATGTCACAAAATTCCTCGGGCTAGATGGAGATAAATTTTTTT	3223
QY	3152	TCTTCTCAGCTTTATGAGGAGAGGAAACTGCTTAGAGTTCAGCTGAACCAACCAAGAAC	3211
Db	3224	TCTTCTCAGCTTTATGAGGAGAGGAAACTGCTTAGAGTTCAGCTGAACCAACCAAGAAC	3283
QY	3212	CTGGCAACATCAGATTTAAGCTTAAGCTTGGGAGCTAACGAGTCTACCTCCCTCTTTGT	3271
Db	3284	CTGGCAACATCAGATTTAAGCTTAAAGTTGGGAGCTAACGAGTCTACCTCCCTCTTTGT	3343
QY	3272	AAATCAAGAATTTGTTTAAATCGGATTCCTCAATCCTTTAAATAAAGATGAACCTTGGTT	3331
Db	3344	AAATCAAGAATTTGTTTAAATCGGATTCCTCAATCCTTTAAATAAAGATGAACCTTGGTT	3403
QY	3332	C 3332	
Db	3404	C 3404	
RESULT 10			
BC042101			
LOCUS			
DEFINITION		3284 bp mRNA linear	PRI 07-OCT-2003
ACCESSION		MG:50665 IMAGE:4400999, complete cds.	
VERSION			
KEYWORDS			
SOURCE		MG:50665 IMAGE:4400999, complete cds.	
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PubMed			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			



REMARK  
COMMENT

USA  
 NIH-MGC Project URL: <http://mgs.nci.nih.gov>  
 Contact: MGC help desk  
 Email: [cgabs@mail.nih.gov](mailto:cgabs@mail.nih.gov)  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 Center, Stanford University School of Medicine, Stanford, CA 94305  
 Web site: <http://www.shgc.stanford.edu>  
 Contact: (Dickson, Mark) [mcdpaxil.stanford.edu](mailto:mcdpaxil.stanford.edu)  
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAX Plate: 88 Row: a Column: 17  
 This clone was selected for full length sequencing because it passed the following selection criteria: GenomeScan gene prediction.

FEATURES  
source

1. 3284  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="MGC:50665 IMAGE:4400399"  
 /tissue\_type="Duodenum, adenocarcinoma"  
 /clone\_lib="NIH MGC\_88"  
 /lab\_host="DH10B"  
 /note="Vector: PCMV-SPORT6"

## gene

1. 3284  
 /gene="DUSP16"  
 /note="synonyms: KIAA1700, MKP-7, MKP7"  
 /db\_xref="LocusID:80824"  
 /db\_xref="MIM:607175"  
 332..2320

## CDS

/codon\_start=1  
 /product="DUSP16 protein"  
 /protein\_id="AAH42101.1"  
 /db\_xref="GI:27469789"  
 /translation="MAHEMIGTQIVTERLVALLESGETKVLILDSPPFVYNTSHILE  
 AININCSKMRRLQDQVILTELIOHSAKHVDIDCSQKVYVDOSSQDVALSLSDC  
 FLTVLLGLKLEKSPVHLLAGFAEFKPGCEKSTLVPTCISQCLPVANTGPT  
 RILPNLYLGCORVNLKMGQNGIYVLANSTCKPDPFIPESHFLRVNDSCEK  
 ILPWLDKSDVRIEKAKSNGVLYHCLAGISRSNTALAIYIKRMDMSLDEAYRVE  
 KEPTISPNFLQLLDYEKKIKNOTGASGPKSLKLLHLEKNEPFPVAVSGGQKSE  
 TPLSPCADSATSAGORVHPASVSVQSLLESPLVQALSGHLHLSADRLSDNK  
 LKRSFLDIKSVSVSAGNMAASLHGFSSSEDALEYKFTLDTNKLKQFSPVQLSE  
 QTPETSPDKEASIPKKIQTARPSDSQSKELSVRTSSGTAQRSLPLHRSQVSD  
 NYHTFLGLSTSQOHLTKAGLKGWHSIDILAPOTS"PSLTSSWYFATESSHFYSA  
 SAIYGGASYSAYSQSLPTCGDOVSVRRRQKPSDRADRSRSHBESPFKQPKRRS  
 COMEFGESIMSENRSRELKGVGSQSFSFGSMELIEVS"  
 380..730

## misc\_feature

/note="RHOD; Region: Rhodanese Homology Domain"  
 /db\_xref="CDD:smart00450"

## misc\_feature

/note="DSPc; Region: Dual specificity phosphatase,  
 catalytic domain. Ser/Thr and Tyr protein phosphatases.  
 The enzyme's tertiary fold is highly similar to that of  
 tyrosine-specific phosphatases, except for a 'recognition'  
 region"  
 /db\_xref="CDD:pfam00782"

## ORIGIN

Query Match 86.7%; Score 2888; DB 9; Length 3284;  
 Best Local Similarity 94.7%; Pred. No. 0;  
 Matches 3092; Conservative 0; Mismatches 0; Indels 174; Gaps 3;  
 231 GAGCAAGGTAAAGATCATGTATCGCTGGCTGCCAAGACATCTTTGTGTGGAA 290  
 1 GAGCAAGGTAAAGATCATGTATCGCTGGCTGCCAAGACATCTTTGTGTGGAA 60

QY 291 TGGTTATTCAGTCATCTCTTTATGAATCAAAATGTGAGGGGCTGCTTTGTGACGGAGTC 350  
 Db 61 TGGTTATTCAGTCATCTCTTTATGAATCAAAATGTGAGGGGCTGCTTTGTGACGGAGTC 120  
 QY 351 CTTTGCAGAGCATCATCAACGGGAAAGAGAAAGACATTCACITTTGGAGGGCTCTTGTCTG 410  
 Db 121 CTTTGCAGAGCATCATCAACGGGAAAGAGAAAGACATTCACITTTGGAGGGCTCTTGTCTG 180  
 QY 411 AAAATGGGTTTAACTCTCTCTTTTGGCAGTCAACACAGCCTGACCTCATACACITTTAGT 470  
 Db 181 AAAATGGGTTTAACTCTCTCTTTTGGCAGTCAACACAGCCTGACCTCATACACITTTAGT 240  
 QY 471 ACATGGAGTGGCTGAGCCTTTTGAGCACACACCATTTACATCATCTGCGCAAAATAAGA 530  
 Db 241 ACATGGAGTGGCTGAGCCTTTTGAGCACACACCATTTACATCATCTGCGCAAAATAAGA 300  
 QY 531 AGGAGTGGGAAAGAGAGACTTATTGTGTCTCATGCCCATTGAGATGATTGGAATCAAAAT 590  
 Db 301 AGGAGTGGGAAAGAGAGACTTATTGTGTCTCATGCCCATTGAGATGATTGGAATCAAAAT 360  
 QY 591 TGTACTGAGAGGTTGGTGGCTCTGCTGGAAGTGGAAACGGAAGAGTGGCTGCTAATTGA 650  
 Db 361 TGTACTGAGAGGTTGGTGGCTCTGCTGGAAGTGGAAACGGAAGAGTGGCTGCTAATTGA 420  
 QY 651 TAGCCGCCCATTTTGTGGAATACAATACATCCACATTTTGGAAAGCCATTAAATATCAACTG 710  
 Db 421 TAGCCGCCCATTTTGTGGAATACAATACATCCACATTTTGGAAAGCCATTAAATATCAACTG 480  
 QY 711 CTCRAAGCTTATGAAGCGAAGTTGCAACAGACAGCAAACTGTTAATTACAGAGCTCATCCA 770  
 Db 481 CTCRAAGCTTATGAAGCGAAGTTGCAACAGACAGCAAACTGTTAATTACAGAGCTCATCCA 540  
 QY 771 GCATTGAGCGAAACATAGAGTTGACATTTGATTCAGTCAAGAGTGTAGTTTACGATCA 830  
 Db 541 GCATTGAGCGAAACATAGAGTTGACATTTGATTCAGTCAAGAGTGTAGTTTACGATCA 600  
 QY 831 AAGTCCCAAGATGTGGCTCTCTCTCTGACACTGTTTCTCACTGACTTCTCTGGGTAA 890  
 Db 601 AAGTCCCAAGATGTGGCTCTCTCTCTGACACTGTTTCTCACTGACTTCTCTGGGTAA 660  
 QY 891 ACTGGAGAAGAGCTTCAACTCTGTTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 926  
 Db 661 ACTGGAGAAGAGCTTCAACTCTGTTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720  
 QY 927 ----- 926  
 Db 721 TTGTTTCCCTGGCTCTGTGGAAGAAATCCACTAGTCCCTACCTGCTGCTGCTGCTGCTGCTGCT 780  
 QY 927 ----- 926  
 Db 781 TTGCTTACCTGTTGCCAACATTGGGCCAACCGGAATTTCTCCCAATCTTTATCTTTGCTG 840  
 QY 927 -----AGGAGCTGATGACGAGATGGGATTGGTATGTTGTTAAA 966  
 Db 841 CAGCGAGATGTCCTCAACAGAGAGCTGATGAGGAGATGGGATTGGTATGTTGTTAAA 900  
 QY 967 TGGCAGCAATACCTGTCCAAAGCCTGACTTTTATCCCGAGTCTCATTTCTCTGCGTGTGCC 1026  
 Db 901 TGGCAGCAATACCTGTCCAAAGCCTGACTTTTATCCCGAGTCTCATTTCTCTGCGTGTGCC 960  
 QY 1027 TGTGAATGACAGCTTTTGTGAGAAATTTTGCCTGCTGTGGCAAAATCAGTAGATTTTAT 1086  
 Db 961 TGTGAATGACAGCTTTTGTGAGAAATTTTGCCTGCTGTGGCAAAATCAGTAGATTTTAT 1020  
 QY 1087 TGAGAAAGCAAAAGCCTCCAATGGATGTGTTTCTAGTGCACCTGTTTAGCTGGGATCTCCCG 1146  
 Db 1021 TGAGAAAGCAAAAGCCTCCAATGGATGTGTTTCTAGTGCACCTGTTTAGCTGGGATCTCCCG 1080  
 QY 1147 CTCGCCACCATCGGTATCGCTATCATCATCAAGAGATGAGCATGCTCTTTAGATGAGC 1206  
 Db 1081 CTCGCCACCATCGGTATCGCTATCATCATCAAGAGATGAGCATGCTCTTTAGATGAGC 1140  
 QY 1207 TTAAGATTTGTAAGAAAGAAAGAACCTTACTATATCTCCAAACTTCAATTTTCTGGGCGCA 1266



Db 1141 TTACAGATTTGTGAAGAAAGAAAGACCTACTATATCTCCAAACTTCAATTTTCTGGGCCA 1200  
QY 1267 ACTCCTGACTATGAGAAAGAAATTAAAGACAGACTGGAGCATCAGGGCCAAAGAGCAA 1326  
Db 1201 ACTCCTGACTATGAGAAAGAAATTAAAGACAGACTGGAGCATCAGGGCCAAAGAGCAA 1260  
QY 1327 ACTAAGCTGTGTGACCTGGAGAGCCAAATGAACCTGTCCCTGCTGTCTCAGAGGGTGG 1386  
Db 1261 ACTAAGCTGTGTGACCTGGAGAGCCAAATGAACCTGTCCCTGCTGTCTCAGAGGGTGG 1320  
QY 1387 ACAGAAAAGAGAGAGCGCCCTCAGTCCACCTGTGCGGACTCTCTACCTCAGAGGCGAGC 1446  
Db 1321 ACAGAAAAGAGAGAGCGCCCTCAGTCCACCTGTGCGGACTCTCTACCTCAGAGGCGAGC 1380  
QY 1447 AGGACAAAGGCCGTGCAATCCGCCAGCGTCCAGCGTCCAGCGTCCAGCGTCCAGCGTCCGCT 1506  
Db 1381 AGGACAAAGGCCGTGCAATCCGCCAGCGTCCAGCGTCCAGCGTCCAGCGTCCAGCGTCCGCT 1431  
QY 1507 GTTAGGAGACGCCGCTGGTACAGCGCTCAGTGGGCTGCACCTGTCCGCGAGCAGCGCT 1566  
Db 1432 GTTAGGAGACGCCGCTGGTACAGCGCTCAGTGGGCTGCACCTGTCCGCGAGCAGCGCT 1491  
QY 1567 GGAAGACAGCAATAAGCTCAAGCGTTCCTTCTCTCTGATATCAAAATCAGTTTCATATTC 1626  
Db 1492 GGAAGACAGCAATAAGCTCAAGCGTTCCTTCTCTCTGATATCAAAATCAGTTTCATATTC 1551  
QY 1627 AGCCAGCATGGCAGCATCCTTACATGGTCTCTCTCATCAGAAGATGCTTTGGAACTACTA 1686  
Db 1552 AGCCAGCATGGCAGCATCCTTACATGGTCTCTCTCATCAGAAGATGCTTTGGAACTACTA 1611  
QY 1687 CAAACCTTCCACTTACTCTGGATGGACCAACAAAGCTATGCCAGTTCCTCCCTGTTCAGGA 1746  
Db 1612 CAAACCTTCCACTTACTCTGGATGGACCAACAAAGCTATGCCAGTTCCTCCCTGTTCAGGA 1671  
QY 1747 ACTATCGGAGCAGACTCCGGAACAGCTCTGATTAAGGAGGAGCCAGCATCCGCAAGAA 1806  
Db 1672 ACTATCGGAGCAGACTCCGGAACAGCTCTGATTAAGGAGGAGCCAGCATCCGCAAGAA 1731  
QY 1807 GCTGCAGAGCGCCAGCGCTTCAGACGCGCAGAGCAGAGCGATTCGATTCGGTTCAGAAACAG 1866  
Db 1732 GCTGCAGAGCGCCAGCGCTTCAGACGCGCAGAGCAGAGCGATTCGATTCGGTTCAGAAACAG 1791  
QY 1867 CAGCAGTGGCACCGCCAGAGTTCCTTTATCTCCACTGCAATGGAAGTGGAGCGTGGGA 1926  
Db 1792 CAGCAGTGGCACCGCCAGAGTTCCTTTATCTCCACTGCAATGGAAGTGGAGCGTGGGA 1851  
QY 1927 GGCAATATCACACAGCTTCCTTTTCGCGCTTTTCACAGCCAGCAGCAGCCTCAGCAA 1986  
Db 1852 GGCAATATCACACAGCTTCCTTTTCGCGCTTTTCACAGCCAGCAGCAGCCTCAGCAA 1911  
QY 1987 GTCTGTGGCCTGGGCTTAAAGGCTGGCACTCGGATATCTGGCGCCCGCCAGCCTCTAC 2046  
Db 1912 GTCTGTGGCCTGGGCTTAAAGGCTGGCACTCGGATATCTGGCGCCCGCCAGCCTCTAC 1971  
QY 2047 CCCTTCCCTCAGCAGCAGCTGTATTTTTCGCAAGAGTCCCTCAGCAGCTTCTACTCTGCGCTC 2106  
Db 1972 CCCTTCCCTCAGCAGCAGCTGTATTTTTCGCAAGAGTCCCTCAGCAGCTTCTACTCTGCGCTC 2031  
QY 2107 AGCCATCTACGGAGGCGAGTCCAGTTCCTCTCTGCTTACAGTGCAGCAGCTGCCCACTTG 2166  
Db 2032 AGCCATCTACGGAGGCGAGTCCAGTTCCTCTCTGCTTACAGTGCAGCAGCTGCCCACTTG 2091  
QY 2167 CGGAGCCAGTCTATTTCTGTGCGAGGCGGAGAGCCAAAGTGCAGAGCTGACTCGCG 2226  
Db 2092 CGGAGCCAGTCTATTTCTGTGCGAGGCGGAGAGCCAAAGTGCAGAGCTGACTCGCG 2151  
QY 2227 GCGAGCTGGCATGAAGAGAGCGCCCTTTGAAAAGCAGTGTAAACGAGAGCTGCCAAAT 2286  
Db 2152 GCGAGCTGGCATGAAGAGAGCGCCCTTTGAAAAGCAGTGTAAACGAGAGCTGCCAAAT 2211  
QY 2287 GGAATTTGGAGAGCATCATGTACAGAGACAGTTCACGGAGAGCTCGGGAAGTGGG 2346

RESULT 11  
BD171157

Db 2212 GGAATTTGGAGAGAGCATCATGTACAGAAACAGGTCAACGGGAAGAGCTGGGAAAAGTGGG 2271  
QY 2347 CAGTCAGTCTAGCTTTTTCGGGAGCATGGAATCATTTAGGTCTCTCTGAGAGAAAGACA 2406  
Db 2272 CAGTCAGTCTAGCTTTTTCGGGAGCATGGAATCATTTAGGTCTCTCTGAGAGAAAGACA 2331  
QY 2407 CTTCTGACTCTTATAGACAAATTTTTTTTCTTCTTCAAAAAAATTCCTGTAAATCTG 2466  
Db 2332 CTTCTGACTCTTATAGACAAATTTTTTTTCTTCTTCAAAAAAATTCCTGTAAATCTG 2391  
QY 2467 AAATATATATGTACATATATATTTTTTGGAAAATGGAGCTATGTTGTAAGAGCAA 2526  
Db 2392 AAATATATATGTACATATATATTTTTTGGAAAATGGAGCTATGTTGTAAGAGCAA 2451  
QY 2527 CAGTGTGATCAACCCAGTGTACTCTCTTAAACATCTGCAATTTGAGAGATCAGCTAATAC 2586  
Db 2452 CAGTGTGATCAACCCAGTGTACTCTCTTAAACATCTGCAATTTGAGAGATCAGCTAATAC 2511  
QY 2587 TTCTCTCAACAAAATGGAAGGCGAGATGTAGATCCCTCTAGAGCGGAGGAAAACCAT 2646  
Db 2512 TTCTCTCAACAAAATGGAAGGCGAGATGTAGATCCCTCTAGAGCGGAGGAAAACCAT 2571  
QY 2647 TTATTTCAAGTCAATTAACATCTCTTCTTCTTAAAAAGCAAGTGTCTTTGGTGTGGA 2706  
Db 2572 TTATTTCAAGTCAATTAACATCTCTTCTTCTTAAAAAGCAAGTGTCTTTGGTGTGGA 2631  
QY 2707 GGACAAAATCCCTTACCAATTTTCCAGTGTGTCTAATAAGAGATCTCAATATATAGTCTT 2766  
Db 2632 GGACAAAATCCCTTACCAATTTTCCAGTGTGTCTAATAAGAGATCTCAATATATAGTCTT 2690  
QY 2767 TGTCGGGACCTTCCATAGTACACCTTAGCGCTGAGACTGAGCCAGCTTGGGGGTGAGGT 2826  
Db 2691 TGTCGGGACCTTCCATAGTACACCTTAGCGCTGAGACTGAGCCAGCTTGGGGGTGAGGT 2750  
QY 2827 AGGTAGACCTTGTAGGAGCAGAGCCTAGTGTGTAATCCAGAGAAATGATCTTATCCAA 2886  
Db 2751 AGGTAGACCTTGTAGGAGCAGAGCCTAGTGTGTAATCCAGAGAAATGATCTTATCCAA 2810  
QY 2887 AGCTCATTTCAAAACCCAGCTCACCTTGACAGCGAGGAGCACAGAGCATCACTCTGCTGG 2946  
Db 2811 AGCTCATTTCAAAACCCAGCTCACCTTGACAGCGAGGAGCACAGAGCATCACTCTGCTGG 2870  
QY 2947 ACGGACCAATTAGGGGCTTGCCAAAGTCTACCTTAGAGCAAAACCCAGTACCTCAGACAGG 3006  
Db 2871 ACGGACCAATTAGGGGCTTGCCAAAGTCTACCTTAGAGCAAAACCCAGTACCTCAGACAGG 2930  
QY 3007 AAAGTCGGGCTTTGACCACTACCATATCTGTAGCCCATTTCTAGGCAATTTGGAATAG 3066  
Db 2931 AAAGTCGGGCTTTGACCACTACCATATCTGTAGCCCATTTCTAGGCAATTTGGAATAG 2990  
QY 3067 GTAGGTAGCTAGTACACCTTTTCAGACCAATTCAAACTGTCTATGCAAAATTTCCGCTG 3126  
Db 2991 GTAGGTAGCTAGTACACCTTTTCAGACCAATTCAAACTGTCTATGCAAAATTTCCGCTG 3050  
QY 3127 GGCCTAGATGGAGATAATTTTTTTTTTCTTCTCAGCTTTATGAAGAGAGGAAACTGTCT 3186  
Db 3051 GGCCTAGATGGAGATAATTTTTTTTTTCTTCTCAGCTTTATGAAGAGAGGAAACTGTCT 3110  
QY 3187 AGGATTCAGCTGAAACCCAGGAACTTCGCAACATCACGATTTAAGCTTAAGCTTGGGAGG 3246  
Db 3111 AGGATTCAGCTGAAACCCAGGAACTTCGCAACATCACGATTTAAGCTTAAGCTTGGGAGG 3170  
QY 3247 CTAAACAGTCTACCTCCCTCTTTGTAAATCAAGAGATTTGTTAAATGGGATTTGCTAATC 3306  
Db 3171 CTAAACAGTCTACCTCCCTCTTTGTAAATCAAGAGATTTGTTAAATGGGATTTGCTAATC 3230  
QY 3307 CTTTAAATAAGATGAACCTTGGTTTC 3332  
Db 3231 CTTTAAATAAGATGAACCTTGGTTTC 3256

LOCUS	BD171157	Novel gene and protein encoded thereby.	4790 bp	DNA	linear	PAT 17-JAN-2003
DEFINITION	BD171157	Novel gene and protein encoded thereby.	4790 bp	DNA	linear	PAT 17-JAN-2003
ACCESSION	BD171157	Novel gene and protein encoded thereby.	4790 bp	DNA	linear	PAT 17-JAN-2003
VERSION	BD171157.1	GI:27876969				
KEYWORDS	WO 02052005-A/13.					
SOURCE	Homo sapiens (human)					
ORGANISM	Homo sapiens					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
AUTHORS	Ohara, O., Nagase, T. and Nakajima, D.					
TITLE	Novel gene and protein encoded thereby					
JOURNAL	Patent: WO 02052005-A 13 04-JUL-2002; KAZUSA DNA RESEARCH INSTITUTE, OSAMU OHARA, TAKAHIRO NAGASE, DAISUKE NAKAJIMA					
COMMENT	OS Homo sapiens (human) PN WO 02052005-A/13 PD 04-JUL-2002 PF 20-DEC-2001 WO 2001JP011217 PR 22-DEC-2000 JP 00P 389742 PI OSAMU OHARA, TAKAHIRO NAGASE, DAISUKE NAKAJIMA PC C12N15/12, C07K14/47 CC Novel gene and protein encoded thereby FH Key Location/Qualifiers FT CDS Location/Qualifiers 1 4790 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606"					
FEATURES	source					
ORIGIN	Query Match 83.0%; Score 2766.4; DB 6; Length 4790; Best Local Similarity 94.7%; Pred. No. 0; Matches 2952; Conservative 0; Mismatches 1; Indels 165; Gaps 2;					
QY	379	GAAGAAGACATTCACCTTGAGGGCTCTTCCTGAAATGGGTTAACTCTCCTTTGGCCAG	438			
DB	1	GAAGAAGACATTCACCTTGAGGGCTCTTCCTGAAATGGGTTAACTCTCCTTTGGCCAG	60			
QY	439	TCACCAACCGCTGACCTCATACATCTTTAGTACATGAGTGGCTGAGCTTTGAGCAC	498			
DB	61	TCACCAACCGCTGACCTCATACATCTTTAGTACATGAGTGGCTGAGCTTTGAGCAC	120			
QY	499	ACCACCATTAACATCATCGTGGCAAAATTAAGAAGGAGGTGGGAAAGAGGACTTATTGTT	558			
DB	121	ACCACCATTAACATCATCGTGGCAAAATTAAGAAGGAGGTGGGAAAGAGGACTTATTGTT	180			
QY	559	GTATGCCCCATCAGATGATTGGAACCTCAATTTGTTACTGAGAGGTGGTGGCTCTGCTG	618			
DB	181	GTATGCCCCATCAGATGATTGGAACCTCAATTTGTTACTGAGAGGTGGTGGCTCTGCTG	240			
QY	619	GAAGTGGAAAGCAAAAGTGCCTGCTAAATGATAGCGGCCCATTTGTGGAATACAATACA	678			
DB	241	GAAGTGGAAAGCAAAAGTGCCTGCTAAATGATAGCGGCCCATTTGTGGAATACAATACA	300			
QY	679	TCCCAATTTTGAAGCCATTAAATATCACTGCTCCAGCTTTAGAGGAGGTTGCA	738			
DB	301	TCCCAATTTTGAAGCCATTAAATATCACTGCTCCAGCTTTAGAGGAGGTTGCA	360			
QY	739	CAGGACAAAGTGTAAATACAGAGCTCATCCAGCTTTAGAGGAGGTTGCA	798			
DB	361	CAGGACAAAGTGTAAATACAGAGCTCATCCAGCTTTAGAGGAGGTTGCA	420			
QY	799	GATTGCACTCAGAGGTTGTAGTTTACGATCAAGCTCCCAAGATGTTGCTCTCTCTCT	858			
DB	421	GATTGCACTCAGAGGTTGTAGTTTACGATCAAGCTCCCAAGATGTTGCTCTCTCTCT	480			
QY	859	TCAGACTGTTTTCTCACTGTACTCTCTGGTAACTGGAGAGGCTTCACTCTCTCTCAC	918			
DB	481	TCAGACTGTTTTCTCACTGTACTCTCTGGTAACTGGAGAGGCTTCACTCTCTCTCAC	540			
QY	919	CTGCTTGC-----	926			

Db	1621	CAGACGAAGCGAATTGCATTCGGTCAAGAACGACGACGAGTGCGGACCGCCCGACGAGTCCCTTT	1680
Qy	1895	TTATCTCCACTGCAATCGAAAGTGGGAGCGTGGAGGACAAATTACCAACACGACGCTTCCCTTTTC	1954
Db	1681	TTATCTCCACTGCAATCGAAAGTGGGAGCGTGGAGGACAAATTACCAACACGACGCTTCCCTTTTC	1740
Qy	1955	GGCCTTTTCCACGACGACGACGACCTCAAGAGTCTGCTGGCTTGGGCTTTAGGGCTCG	2014
Db	1741	GGCCTTTTCCACGACGACGACGACCTCAAGAGTCTGCTGGCTTGGGCTTTAGGGCTCG	1800
Qy	2015	CACCTCGGATATCTTTGGGCCCGCCAGACCTCTACCCCTTCCCTGACCCAGCAGCTGGTATTTTT	2074
Db	1801	CACCTCGGATATCTTTGGGCCCGCCAGACCTCTACCCCTTCCCTGACCCAGCAGCTGGTATTTTT	1860
Qy	2075	GCCACAGAGTCCCTCACACTTCTACTCTGGCTCAGCCATCTACGAGGAGGAGTCCGAGTTAC	2134
Db	1861	GCCACAGAGTCCCTCACACTTCTACTCTGGCTCAGCCATCTACGAGGAGGAGTCCGAGTTAC	1920
Qy	2135	TCGTGCTACAGCTGCAGCAGCTGCCACTTGGGAGAGACCAGATCTATCTGTGCGCAGG	2199
Db	1921	TCGTGCTACAGCTGCAGCAGCTGCCACTTGGGAGAGACCAGATCTATCTGTGCGCAGG	1980
Qy	2195	GGCGAGAAGCCAAAGTGACAGAGTGACTCGCGCGCGAGCTGGCATGAAGAGAGCCCTCTTT	2254
Db	1981	GGCGAGAAGCCAAAGTGACAGAGTGACTCGCGCGCGAGCTGGCATGAAGAGAGCCCTCTTT	2040
Qy	2255	GAAAGCAGTTTTAAACCCAGAGCTGCCAAATGGAATTTGGAGAGAGCATCATGTACAGAG	2314
Db	2041	GAAAGCAGTTTTAAACCCAGAGCTGCCAAATGGAATTTGGAGAGAGCATCATGTACAGAG	2100
Qy	2315	AACAGGTCACGGGGAAGAGCTGGGGAAGTGGGCAGTCTAGCTTTTGGGCGAGCATG	2374
Db	2101	AACAGGTCACGGGGAAGAGCTGGGGAAGTGGGCAGTCTAGCTTTTGGGCGAGCATG	2160
Qy	2375	GAAATCATTTAGGTTCTCTGAGAGAAAGACACTGTGACTTCTATAGACAAATTTTTTTTT	2434
Db	2161	GAAATCATTTAGGTTCTCTGAGAGAAAGACACTGTGACTTCTATAGACAAATTTTTTTTT	2220
Qy	2435	TCCTGTTCCACAAAAAATTCCTTGTAATCTGGAATATATATATGACATACATATATAT	2494
Db	2221	TCCTGTTCCACAAAAAATTCCTTGTAATCTGGAATATATATATGACATACATATATAT	2280
Qy	2495	TTTTTGGAAAATGGAGCTATGGTGTAAAGCAACAGGTGGATCAACCCAGTTGTTACTCTC	2554
Db	2281	TTTTTGGAAAATGGAGCTATGGTGTAAAGCAACAGGTGGATCAACCCAGTTGTTACTCTC	2340
Qy	2555	TTAAATCATGCTATTTGAGAGATCAGCTAATACTTCTCTCAACAAAAATGGAAGGCGAGAT	2614
Db	2341	TTAAATCATGCTATTTGAGAGATCAGCTAATACTTCTCTCAACAAAAATGGAAGGCGAGAT	2400
Qy	2615	GCTAGAAATCCCCCTTAGACGGAGGAAAACCATTTTATTCAGTGAATTAACATCCTCTTG	2674
Db	2401	GCTAGGATCCCCCTTAGACGGAGGAAAACCATTTTATTCAGTGAATTAACATCCTCTTG	2460
Qy	2675	TTCTTTAAAAAGCAAGTGTCTTTGGTGTGGAGGACAAAATCCCTACCATTTTCCAGCT	2734
Db	2461	TTCTTTAAAAAGCAAGTGTCTTTGGTGTGGAGGACAAAATCCCTACCATTTT-CACTG	2519
Qy	2735	TGTGCTACTTAAGAGATCTCAAAATTAAGTCTTTGTCCGGAGCCCTTCCCATAGTACACCTTA	2794
Db	2520	TGTGCTACTTAAGAGATCTCAAAATTAAGTCTTTGTCCGGAGCCCTTCCCATAGTACACCTTA	2579
Qy	2795	GGCTGAGACTGAGCCAGCTTTGGGGTTCAGGTAGGTAGACCTGTTTAGGGACAGAGCCTTA	2854
Db	2580	GGCTGAGACTGAGCCAGCTTTGGGGTTCAGGTAGGTAGACCTGTTTAGGGACAGAGCCTTA	2639
Qy	2855	GTGGTAAATCCAAAGAGAAATCATCTTAATCCAAAGCTGATTCACAAACCCACGCTCACCTG	2914
Db	2640	GTGGTAAATCCAAAGAGAAATCATCTTAATCCAAAGCTGATTCACAAACCCACGCTCACCTG	2699
Qy	2915	ACAGCCGAGGACACGAGCATCACTCTGCTGGACGGACCAATTTAGGGGCTTTGCCAAGGTC	2974
Db	2700	ACAGCCGAGGACACGAGCATCACTCTGCTGGACGGACCAATTTAGGGGCTTTGCCAAGGTC	2759

QY	2975	TACCTTAGACAAACCCAGTACCTCAGACAGAAAGTCGGGGCTTTGACACACTACCATAT	3033
DB	2760	TACCTTAGACAAACCCAGTACCTCAGACAGAAAGTCGGGGCTTTGACACACTACCATAT	2819
QY	3035	CTGCTAGCCCATTTTCTAGGCATTTGTGAATAGTAGTAGGTCAGTCTCCTTTTTCAGACC	3094
DB	2820	CTGCTAGCCCATTTTCTAGGCATTTGTGAATAGTAGTAGGTCAGTCTCCTTTTTCAGACC	2879
QY	3095	AATTCAAACTGTCATGTGCACAAAATTCCTGGGCTAGATGGAGATAATTTTTTTTCT	3154
DB	2880	AATTCAAACTGTCATGTGCACAAAATTCCTGGGCTAGATGGAGATAATTTTTTTTCT	2939
QY	3155	TCTCAGCTTTATGAAGAGAGGGAACCTCTCTAGGATTCAGCTGAAACCAACGAGAACCTG	3214
DB	2940	TCTCAGCTTTATGAAGAGAGGGAACCTCTCTAGGATTCAGCTGAAACCAACGAGAACCTG	2999
QY	3215	GCAACATCAGATTTAAGCTTAAGTTGGAGGCTTAACGAGTCTACCTCCCTCTTTGTAAA	3274
DB	3000	GCAACATCAGATTTAAGCTTAAGTTGGAGGCTTAACGAGTCTACCTCCCTCTTTGTAAA	3059
QY	3275	TCAAAGAATTGTTTAAATGGGATTCCTCAATCCTTTAAATAAGATGAACCTTGTTTTC	3332
DB	3060	TCAAAGAATTGTTTAAATGGGATTCCTCAATCCTTTAAATAAGATGAACCTTGTTTTC	3117
RESULT 12			
BD183422			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
1 (bases 1 to 4790)			
Ohara,O., Nagase,T. and Nakajima,D.			
Novel genes and proteins encoded by the genes			
Patient: JP 2002345492-A 135 03-DEC-2002;			
KAZUSA DNA RESEARCH INSTITUTE			
OS Homo sapiens (human)			
PN JP 2002345492-A/135			
PD 03-DEC-2002			
PF 26-FEB-2002 JP 2002049009			
PI OSAMO OHARA, TAKAHIRO NAGASE, DAISUKE NAKAJIMA			
PC C12N15/09,C07K14/47//A61K31/711.A61K38/00,A61K48/00,A61P25/00,			
A61P25/14,			
PC A61P25/18,A61P35/00,C12N15/00,A61K37/02			
CC Novel genes and proteins encoded by the genes FH Key			
Location/Qualifiers			
FT CDS (184)..(2178).			
Location/Qualifiers			
1..4790			
/organism="Homo sapiens"			
/mol_type="genomic DNA"			
/db_xref="taxon:9606"			
FEATURES			
source			
Query Match 83.08; Score 2766.4; DB 6; Length 4790;			
Best Local Similarity 94.7%; Pred. No. 0;			
Matches 2952; Conservative 0; Mismatches 1; Indels 165; Gaps 2;			
ORIGIN			
379 GAAAGAGACATTCACTTGGAGGCTCTTCTGAAATGGTTTAACTCTCCTTTTGCAG 438			
1 GAAAGAGACATTCACTTGGAGGCTCTTCTGAAATGGTTTAACTCTCCTTTTGCAG 60			
439 TCACCACAGGCTGACCTCATACATTTAGTACATGGAGTGGCTGAGCTTTGAGCAC 498			
61 TCACCACAGGCTGACCTCATACATTTAGTACATGGAGTGGCTGAGCTTTGAGCAC 120			
499 ACCACATTATCATCTCGTCGCAAAATTAAGAGAGGAGTGGGAAAGAGGACTTATTGTT 558			

Db 121 ACCACCAATTACATCATCGTGCAAAATTAAGAGAGGAGGTGGGAAAGAGACTTAATGTT 180  
Qy 559 GTCATGGCCCATGAGATGAGATTGGAACCTCAAATGTTACTGAGAGGTTGGTGGCTCTGCTG 618  
Db 181 GTCATGGCCCATGAGATGAGATTGGAACCTCAAATGTTACTGAGAGGTTGGTGGCTCTGCTG 240  
Qy 619 GAAAGTGGAAAGGAAAGTGGCTGCTAAATGATAGCGGCCCATTTGTGGAATACAAATACA 678  
Db 241 GAAAGTGGAAAGGAAAGTGGCTGCTAAATGATAGCGGCCCATTTGTGGAATACAAATACA 300  
Qy 679 TCCACACATTTTGGAAAGCATTAAATATCAACTGCTCCAAAGCTTATGAAGCGAAGTTGCA 738  
Db 301 TCCACACATTTTGGAAAGCATTAAATATCAACTGCTCCAAAGCTTATGAAGCGAAGTTGCA 360  
Qy 739 CAGGACAAAGTGTAAATACAGAGCTCATCCAGACTTCAGCGAAATCAATAGGTTGACATT 798  
Db 361 CAGGACAAAGTGTAAATACAGAGCTCATCCAGACTTCAGCGAAATCAATAGGTTGACATT 420  
Qy 799 GATTGCGAGTCAGAGGTTGTAGTTTACGATCAAGCTCCCAAGATGTTGCCCTCTCTCT 858  
Db 421 GATTGCGAGTCAGAGGTTGTAGTTTACGATCAAGCTCCCAAGATGTTGCCCTCTCTCT 480  
Qy 859 TCAGACTGTTTCTCACTGACTCTCTGGGTAATCTGGAGAGAGCTTCAACTCTGTTTCA 918  
Db 481 TCAGACTGTTTCTCACTGACTCTCTGGGTAATCTGGAGAGAGCTTCAACTCTGTTTCA 540  
Qy 919 CTGCTTGC----- 926  
Db 541 CTGCTTGCAGGTGGTGTGAGTTCCTCGTGTGTTTCCCTGGGCTCTGTGAAGGAAAA 600  
Qy 927 ----- 926  
Db 601 TCCACTCTAGTCCCTACCTGCAATTCAGSCCTTGCTTAACCTGTGCGCAACATTGGGCA 660  
Qy 927 -----AGAGCTG 934  
Db 661 ACCCGAATTTCCCAATCTTTATCTTGGCTGCGAGAGATGCTCTCAACAGGAGCTG 720  
Qy 935 ATGAGAGAGATGGATGGTGTATGCTTAAATGCGAGCAATACCTGTCCAAAGCTGAC 994  
Db 721 ATGAGAGAGATGGATGGTGTATGCTTAAATGCGAGCAATACCTGTCCAAAGCTGAC 780  
Qy 995 TTTATCCCGAGTCTCATTTCTCGTGGTGGCTGTGAATGACAGCTTTTGTGAGAAAT 1054  
Db 781 TTTATCCCGAGTCTCATTTCTCGTGGTGGCTGTGAATGACAGCTTTTGTGAGAAAT 840  
Qy 1055 TTGCGGTGGTGGACAAATCAGTAGATTTTCAATTGAGAAAGCAAAAGCCCTCAATGGATGT 1114  
Db 841 TTGCGGTGGTGGACAAATCAGTAGATTTTCAATTGAGAAAGCAAAAGCCCTCAATGGATGT 900  
Qy 1115 GTTCTAGTGCACTGTTTAGCTGGGATCTCCCGTCCCGCCACCATCTGCTTACATC 1174  
Db 901 GTTCTAGTGCACTGTTTAGCTGGGATCTCCCGTCCCGCCACCATCTGCTTACATC 960  
Qy 1175 ATGAAGAGATGGACATGCTTTTAGATGAGCTTACAGATTTGTGAAGAAAAAGACCT 1234  
Db 961 ATGAAGAGATGGACATGCTTTTAGATGAGCTTACAGATTTGTGAAGAAAAAGACCT 1020  
Qy 1235 ACTATATCTCCAAACTTCAATTTTCTGGGCCAACTCTCTGCACTATCAGAAAGAAATTAA 1294  
Db 1021 ACTATATCTCCAAACTTCAATTTTCTGGGCCAACTCTCTGCACTATCAGAAAGAAATTAA 1080  
Qy 1295 AACAGACTGGAGCATCAGGGCCAAAGACAACTCAAGCTGTGCACTGGAGAGGCA 1354  
Db 1081 AACAGACTGGAGCATCAGGGCCAAAGACAACTCAAGCTGTGCACTGGAGAGGCA 1140  
Qy 1355 AATGAACTGTCCCTGTGCTCAGAGGGTGGACAGAAAGCGAGAGCGCCCTCAGTCCA 1414  
Db 1141 AATGAACTGTCCCTGTGCTCAGAGGGTGGACAGAAAGCGAGAGCGCCCTCAGTCCA 1200  
Qy 1415 CCCTGTGCCGACTCTGCTACCTCAGAGGCGAGGAGCAAAAGGCCCGGTGATCCCGCCAGC 1474

Db 1201 CCCTGTGCCGACTCTCTGCTACCTCAGAGGCGAGCGAGCAAAAGGCCCGTGTGATCCCGCCAGC 1260  
Qy 1475 GTGCCACGCGTGGCCACGCTGCGAGCGTGGTGTAGAGGACAGCGCGTGGTACAGGCG 1534  
Db 1261 GTGCCACGCGTGGCCACGCTGCGAGCGTGGTGTAGAGGACAGCGCGTGGTACAGGCG 1320  
Qy 1535 CTGAGTGGGCTGCACCTGTCCGAGACAGGCTGGGAAAGACAGCAATAGCTCAAGCGTTCC 1594  
Db 1321 CTGAGTGGGCTGCACCTGTCCGAGACAGGCTGGGAAAGACAGCAATAGCTCAAGCGTTCC 1380  
Qy 1595 TTCTCTCTGATATCAAACTCAGTTTCATATTCAGCCAGCATGGCAGCATCTTATCATGGC 1654  
Db 1381 TTCTCTCTGATATCAAACTCAGTTTCATATTCAGCCAGCATGGCAGCATCTTATCATGGC 1440  
Qy 1655 TTCTCTCTCATCAGAGATGCTTTGGAAATCTCAAACTTCCACTCTCTGATGGGACC 1714  
Db 1441 TTCTCTCTCATCAGAGATGCTTTGGAAATCTCAAACTTCCACTCTCTGATGGGACC 1500  
Qy 1715 AACAACTATGCCAGTCTTCCCTGTTTCAAGGAATCTCGAGCAGACTCCCGAAACCACT 1774  
Db 1501 AACAACTATGCCAGTCTTCCCTGTTTCAAGGAATCTCGAGCAGACTCCCGAAACCACT 1560  
Qy 1775 CTTGATAAGGAGAAAGCCAGCATCCCAAGAGCTGCAGACCGCCAGGCTTTCAGACAGC 1834  
Db 1561 CTTGATAAGGAGAAAGCCAGCATCCCAAGAGCTGCAGACCGCCAGGCTTTCAGACAGC 1620  
Qy 1835 CAGAGCAAGCGATTGCTTCGGTTCAGAACCCAGCAGTGCGCACCGCCAGAGGTCCTT 1894  
Db 1621 CAGAGCAAGCGATTGCTTCGGTTCAGAACCCAGCAGTGCGCACCGCCAGAGGTCCTT 1680  
Qy 1895 TTATCTCCACTGATCAAGAGTGGGAGCGTGGAGGACAAATACCAACCAAGCTTCCCTTTTC 1954  
Db 1681 TTATCTCCACTGATCAAGAGTGGGAGCGTGGAGGACAAATACCAACCAAGCTTCCCTTTTC 1740  
Qy 1955 GGCTTTTCCACAGCCAGCAGCAGCACTCAGAAAGTCTGCTGGCTGGGCTTAAAGGGCTGG 2014  
Db 1741 GGCTTTTCCACAGCCAGCAGCAGCACTCAGAAAGTCTGCTGGCTGGGCTTAAAGGGCTGG 1800  
Qy 2015 CACTCGGATATCTTGGCCCGCCAGACCTCTACCCCTTCCCTGACAGCAGCTGGTATTTT 2074  
Db 1801 CACTCGGATATCTTGGCCCGCCAGACCTCTACCCCTTCCCTGACAGCAGCTGGTATTTT 1860  
Qy 2075 GCACAGAGTCTCACTCTTACTCTGCTCAGCCATCTACGAGGCGAGTGCAGGTTAC 2134  
Db 1861 GCACAGAGTCTCACTCTTACTCTGCTCAGCCATCTACGAGGCGAGTGCAGGTTAC 1920  
Qy 2135 TCTGCTTACAGCTGCGAGCGAGCTGCCACTTCCGAGAGCCAAAGTCTTCTGTCGCGAG 2194  
Db 1921 TCTGCTTACAGCTGCGAGCGAGCTGCCACTTCCGAGAGCCAAAGTCTTCTGTCGCGAG 1980  
Qy 2195 CGGAGAGCCAAAGTGCAGAGCTGACTCGCGCGGAGCTGGCATGAAGAGAGCCCCCTTT 2254  
Db 1981 CGGAGAGCCAAAGTGCAGAGCTGACTCGCGCGGAGCTGGCATGAAGAGAGCCCCCTTT 2040  
Qy 2255 GAAAGCAGTTTAAACGCGAGAGCTGCCAAATGGAATTTGGAGAGGATCATGTGCAGG 2314  
Db 2041 GAAAGCAGTTTAAACGCGAGAGCTGCCAAATGGAATTTGGAGAGGATCATGTGCAGG 2100  
Qy 2315 AACAGGTCACGGGAAAGCTGGGAAAGTGGGCACTCAGTCTAGCTTTTTCGGGCGAGCATG 2374  
Db 2101 AACAGGTCACGGGAAAGCTGGGAAAGTGGGCACTCAGTCTAGCTTTTTCGGGCGAGCATG 2160  
Qy 2375 GAAATCATTTGAGGTCTCTGAGAGAAAGACACTTGTGAGCTTCTATAGACAATTTTTTTT 2434  
Db 2161 GAAATCATTTGAGGTCTCTGAGAGAAAGACACTTGTGAGCTTCTATAGACAATTTTTTTT 2220  
Qy 2435 TCTTGTTCACAAAAAATTTCCCTGTAAATCTCAAAATATATATGTATACATACATATAT 2494  
Db 2221 TCTTGTTCACAAAAAATTTCCCTGTAAATCTCAAAATATATATGTATACATACATATAT 2280  
Qy 2495 TTTTGAATAATGAGCTATGGTGTAAAGCAACAGGTGGATCAACCCAGTGTGTTACTCTC 2554  
Db 2281 TTTTGAATAATGAGCTATGGTGTAAAGCAACAGGTGGATCAACCCAGTGTGTTACTCTC 2340



Db 421 GATTGAGTCAGAGGTTGTAGTTTACGATCAAGCTCCCAAGATGTTGCCCTCTCTCTCT 480  
Qy TCAGACTGTTTTCTCACTGTAATCTTGGGTAAACTGGAGAAGAGCTTCAACTCTGTTCAAC 918  
Db TCAGACTGTTTTCTCACTGTAATCTTGGGTAAACTGGAGAAGAGCTTCAACTCTGTTCAAC 540  
Qy CTGCTTTC----- 926  
Db CTGCTTCGAGGTGGTTGTGTGAGTCTCTGCTGTTTCCCTGGCCCTGCTGTGAAGAA 600  
Qy ----- 926  
Db TCCACTCTAGTCCCTACCTGCAATTTCTCAGCCTTGCTTACCTGTTGCCAACATGCGCCA 660  
Qy -----AGAGCTG 934  
Db ACCCGAATCTTCCCAATCTTATCTTGGCTGCCAGCAGATGTCTCTCAACAGAGAGCTG 720  
Qy ATGCAGCAGAAATGGATTTGTTATGTGTTAAATGCCAGCAATACCTGTCCAAAGCCTGAC 994  
Db ATGCAGCAGAAATGGATTTGTTATGTGTTAAATGCCAGCAATACCTGTCCAAAGCCTGAC 780  
Qy TTTATCCCGAGTCTATTTCTCGGTGCTGCTGTGAATGACAGCTTTTGTGAGAAATTT 1054  
Db TTTATCCCGAGTCTATTTCTCGGTGCTGCTGTGAATGACAGCTTTTGTGAGAAATTT 840  
Qy TTGCGGTGTTGGAACAATCAGTAGATTTTCATTTGAGAAAGCAAGCCCTCCCAATGATGT 1114  
Db TTGCGGTGTTGGAACAATCAGTAGATTTTCATTTGAGAAAGCAAGCCCTCCCAATGATGT 900  
Qy GTTCTAGTGCACTGTTTGTAGCTGGATCTCCGCTCCGCCCAACCTGCTATGCTGCTACATC 1174  
Db GTTCTAGTGCACTGTTTGTAGCTGGATCTCCGCTCCGCCCAACCTGCTATGCTGCTACATC 960  
Qy ATGAAGAGGATGGACATGCTTTTGTAGTAGAGCTTACAGATTTGTGAAAGCAAGAAAGCCT 1234  
Db ATGAAGAGGATGGACATGCTTTTGTAGTAGAGCTTACAGATTTGTGAAAGCAAGAAAGCCT 1020  
Qy ACTATATCTCCAAACTTCAATTTTCTGGGCCAACTCTCTGGACTATGAGAAGAGATTAG 1294  
Db ACTATATCTCCAAACTTCAATTTTCTGGGCCAACTCTCTGGACTATGAGAAGAGATTAG 1080  
Qy AACAGACTGGAGCATCAGGCCCAAGAGCAAACTCAAGCTGCTGCACTGGAGAAGCCA 1354  
Db AACAGACTGGAGCATCAGGCCCAAGAGCAAACTCAAGCTGCTGCACTGGAGAAGCCA 1140  
Qy AATGAACCTGCTGCTGCTCAGAGGGTGGACAGAAAGCAGACGCGCCCTCAGTCCA 1414  
Db AATGAACCTGCTGCTGCTCAGAGGGTGGACAGAAAGCAGACGCGCCCTCAGTCCA 1200  
Qy CCTGTGCGGACTCTGTACTCTCAGAGGCAAGGACCAAGAGCCCGTGTGATCCCGCCAGC 1474  
Db CCTGTGCGGACTCTGTACTCTCAGAGGCAAGGACCAAGAGCCCGTGTGATCCCGCCAGC 1260  
Qy GTGCCAGCGTCCAGCGTGCAGCGCTGCTGTGTAGAGCAGACCGCGTGGTACAGGG 1534  
Db GTGCCAGCGTCCAGCGTGCAGCGCTGCTGTGTAGAGCAGACCGCGTGGTACAGGG 1320  
Qy CTCAGTGGGCTGCACCTGTCCGAGCAGAGCTGGAAGACAAATAGCTCAAGCGTTCC 1594  
Db CTCAGTGGGCTGCACCTGTCCGAGCAGAGCTGGAAGACAAATAGCTCAAGCGTTCC 1380  
Qy TTCTCTCTGGATATCAATCAGTTTTCATATTCAGCCAGATGCGAGCATCTTACATGGC 1654  
Db TTCTCTCTGGATATCAATCAGTTTTCATATTCAGCCAGATGCGAGCATCTTACATGGC 1440  
Qy TTCTCTCTGGATATCAATCAGTTTTCATATTCAGCCAGATGCGAGCATCTTACATGGC 1714  
Db TTCTCTCTCAAGAGATGCTTTTGAATACTTACAACCTTTCCACTACTCTGGATGGGACC 1500  
Qy AACAGCTATCCGAGTCTCCCTGTTTCCAGAACTATCCGAGCAGACTCCGAAACAGT 1774  
Db AACAGCTATCCGAGTCTCCCTGTTTCCAGAACTATCCGAGCAGACTCCGAAACAGT 1560

Qy 1775 CTTGATAAGAGGAAGCCAGCATCCCAAGAGCTGCGAGACCGCCAGGCGCTTTCAGACAGC 1834  
Db CTTGATAAGAGGAAGCCAGCATCCCAAGAGCTGCGAGACCGCCAGGCGCTTTCAGACAGC 1620  
Qy CAGAGCAAGCGATTGCTATTCGGTTCAGAACCCAGCAGCTGGCAACCGCCAGAGGTCCCTT 1894  
Db CAGAGCAAGCGATTGCTATTCGGTTCAGAACCCAGCAGCTGGCAACCGCCAGAGGTCCCTT 1680  
Qy TTATCTCCATCGCATCGAAGTGGAGCGTGGAGGACAAATACCAACAGCTTCCCTTTTC 1954  
Db TTATCTCCATCGCATCGAAGTGGAGCGTGGAGGACAAATACCAACAGCTTCCCTTTTC 1740  
Qy GGCTTTTCCACAGCCAGCAGCAGCTCAGAACTGCTGGCTGGCTGGCTTAAAGGCTGG 2014  
Db GGCTTTTCCACAGCCAGCAGCAGCTCAGAACTGCTGGCTGGCTGGCTTAAAGGCTGG 1800  
Qy CACTCGATATCTTGGCCCCCAGACCTCTACCCCTTCCCTGACCAGAGCTGGTATTTT 2074  
Db CACTCGATATCTTGGCCCCCAGACCTCTACCCCTTCCCTGACCAGAGCTGGTATTTT 1860  
Qy GCCACAGAGTCTCAGCTTCTACTCTGCTCAGCATCTACGGAGGAGTCCAGTTAC 2134  
Db GCCACAGAGTCTCAGCTTCTACTCTGCTCAGCATCTACGGAGGAGTCCAGTTAC 1920  
Qy TCTGCCCTACAGCTGCGAGCAGCTGCCCACTTGGGAGACCAAGTCTATTCTGTGGCAGG 2194  
Db TCTGCCCTACAGCTGCGAGCAGCTGCCCACTTGGGAGACCAAGTCTATTCTGTGGCAGG 1980  
Qy CGCAGAGCCAAAGTGAAGCTGACTCGCGCGGAGCTGSCATGAAGAGAGCCCCCTT 2254  
Db CGCAGAGCCAAAGTGAAGCTGACTCGCGCGGAGCTGSCATGAAGAGAGCCCCCTT 2040  
Qy GAAAAGCAGTTTAAACCGCAGAAAGCTGCCAAATGGAATTTGGAGAGAGCATATGTCAGAG 2314  
Db GAAAAGCAGTTTAAACCGCAGAAAGCTGCCAAATGGAATTTGGAGAGAGCATATGTCAGAG 2100  
Qy AACAGGTACCGGAAAGCTGGGAAAGTGGGAGTCACTAGCTTTTCGGGCGAGCATG 2374  
Db AACAGGTACCGGAAAGCTGGGAAAGTGGGAGTCACTAGCTTTTCGGGCGAGCATG 2160  
Qy GAAATCAATGAGTCTCCTGAGAAGAAAGACACTTGTGACTTCTATAGACAATTTTTT 2434  
Db GAAATCAATGAGTCTCCTGAGAAGAAAGACACTTGTGACTTCTATAGACAATTTTTT 2220  
Qy TCTTGTTCACAAAATAATCCCTGTAAATCTGAAATATATATATATATATATATATAT 2494  
Db TCTTGTTCACAAAATAATCCCTGTAAATCTGAAATATATATATATATATATATATATAT 2280  
Qy TTTTGGAAATGGAGCTATGTTAAAGCAAGCTGGATCAACCCAGTTCCTTCTCTC 2554  
Db TTTTGGAAATGGAGCTATGTTAAAGCAAGCTGGATCAACCCAGTTCCTTCTCTC 2340  
Qy TTAACATCTGCAATTTGAGAGATCAGCTAATCTTCTCAACAAAATGGAAGGCGAGAT 2614  
Db TTAACATCTGCAATTTGAGAGATCAGCTAATCTTCTCAACAAAATGGAAGGCGAGAT 2400  
Qy GCTAGAAATCCCGCTTAGACGAGGAAAACATTTTATTCAGTGAATTAACATCCTCTG 2674  
Db GCTAGAAATCCCGCTTAGACGAGGAAAACATTTTATTCAGTGAATTAACATCCTCTG 2460  
Qy TTCTTAAAAAGCAAGTGTCTTGTGTTGGAGGACAAAATCCCTTACCATTTTCCAGCT 2734  
Db TTCTTAAAAAGCAAGTGTCTTGTGTTGGAGGACAAAATCCCTTACCATTTT-CAOCT 2519  
Qy TGTGCTACTAAGAGATCTCAATATTTAGTCTTTTTCGGAGCCCTTCCATAGTACACCTTA 2794  
Db TGTGCTACTAAGAGATCTCAATATTTAGTCTTTTTCGGAGCCCTTCCATAGTACACCTTA 2579  
Qy GCGCTGAGACTGAGCCAGCTTTGGGGGTGAGGTAGGTAGAGCCCTGTTAGGAGCAGAGCCTTA 2854  
Db GCGCTGAGACTGAGCCAGCTTTGGGGGTGAGGTAGGTAGAGCCCTGTTAGGAGCAGAGCCTTA 2639





Qy	1281	AGAAGAAGATTAAACAACCACTGGAGCATCAGGGCCAAAGACAAACTCAAGCTCTGCTC	1340
Db	1042	AGAGAAGATTAAACAACCACTGGAGCATCAGGGCCAAAGACAAACTCAAGCTCTGCTC	1101
Qy	1341	ACCTGGAGAAGCCAAATGAACCTGTCTCCTGCTGTCTCAGAGGGTGGACAGAAAGCGAGA	1400
Db	1102	ACCTGGAGAAGCCAAATGAACCTGTCTCCTGCTGTCTCAGAGGGTGGACAGAAAGCGAGA	1161
Qy	1401	CGCCCTCAGTCCACCCCTGTGCGCAGCTCTGCTACCTCAGAGGCAGCAGGACAAAGGCCCG	1460
Db	1162	CGCCCTCAGTCCACCCCTGTGCGCAGCTCTGCTACCTCAGAGGCAGCAGGACAAAGGCCCG	1221
Qy	1461	TGCATCCCGCCAGCGTCCAGCGTGCCAGCGTCAGCGTCGCTGTTTAGAGGACAGCC	1520
Db	1222	TGCATCCCGCCAGCGTCCAGCGTGCCAGCGTCAGCGTCGCTGTTTAGAGGACAGCC	1281
Qy	1521	CGTGTGTACAGCGCTCAGTGGGCTGCACCTGTCCGACAGCGCTGGAAGACAGCAATA	1580
Db	1282	CGCTGTGTACAGCGCTCAGTGGGCTGCACCTGTCCGACAGCGCTGGAAGACAGCAATA	1341
Qy	1581	AGCTCAAGCGTTCCTTCTCTCTGGATCAAAATCAGTTTCATATTCAGCCAGCATCGCAG	1640
Db	1342	AGCTCAAGCGTTCCTTCTCTCTGGATCAAAATCAGTTTCATATTCAGCCAGCATCGCAG	1401
Qy	1641	CATCCTTACATGGCTTCCTCATCAGAAGATGCTTTGGAATACTACAAACCTTCCACTA	1700
Db	1402	CATCCTTACATGGCTTCCTCATCAGAAGATGCTTTGGAATACTACAAACCTTCCACTA	1461
Qy	1701	CTCTGGATGGACCAACAAGCTATGCCAGTTCCTCCCTGTTCTAGGAACTATCGAGCAGA	1760
Db	1462	CTCTGGATGGACCAACAAGCTATGCCAGTTCCTCCCTGTTCTAGGAACTATCGAGCAGA	1521
Qy	1761	CTCCGAAACCCAGTCCCTGATTAAGAGAGAACCCAGCATCCCACAAAGAGCTGCAGCCGCCA	1820
Db	1522	CTCCGAAACCCAGTCCCTGATTAAGAGAGAACCCAGCATCCCACAAAGAGCTGCAGCCGCCA	1581
Qy	1821	GGCCTTCAGACGCCAGCAGCAAGCGATTGCTCGGTTCAGAAACAGCAGCAGTGGCACCG	1880
Db	1582	GGCCTTCAGACGCCAGCAGCAAGCGATTGCTCGGTTCAGAAACAGCAGCAGTGGCACCG	1641
Qy	1881	CCCAGAGGTCCTTTTATCTCCACTGCATCGAAGTGGAGCGTGGAGGACAAATTACCACA	1940
Db	1642	CCCAGAGGTCCTTTTATCTCCACTGCATCGAAGTGGAGCGTGGAGGACAAATTACCACA	1701
Qy	1941	CGAGCTTCCTTTTCGGCCTTTCCACAGCAGCAGACACCTCAGCAAGTCTGTCTGGCCTGG	2000
Db	1702	CGAGCTTCCTTTTCGGCCTTTCCACAGCAGCAGACACCTCAGCAAGTCTGTCTGGCCTGG	1761
Qy	2001	GCCTTAAAGGCTGGCACTCGGATATCTTGCGCCCCACAGACCTCTACCCCTTCCTTCACCA	2060
Db	1762	GCCTTAAAGGCTGGCACTCGGATATCTTGCGCCCCACAGACCTCTACCCCTTCCTTCACCA	1821
Qy	2061	GCAGCTGGTATTTTGCCACAGAGTCCTCAGACTTCCTACTCTGTGCTCAGCCATCTACGGAG	2120
Db	1822	GCAGCTGGTATTTTGCCACAGAGTCCTCAGACTTCCTACTCTGTGCTCAGCCATCTACGGAG	1881
Qy	2121	GCAGTGCAGTTACTCTGTGCTCAGCTGCAGCAGCTGCCACTTGGGAGACCAAGTCT	2180
Db	1882	GCAGTGCAGTTACTCTGTGCTCAGCTGCAGCAGCTGCCACTTGGGAGACCAAGTCT	1941
Qy	2181	ATTCCTGTGCGAGCGCGAGAACCAAGTGCAGAGCTGACTCTCGCGCGGAGCTGGCATG	2240
Db	1942	ATTCCTGTGCGAGCGCGAGAACCAAGTGCAGAGCTGACTCTCGCGCGGAGCTGGCATG	2001
Qy	2241	AAGAGAGCCCTTTTGAAGAAGCAGTTTAAACGCGAGAAGCTGCCAAATGGAATTTGGAGAGA	2300
Db	2002	AAGAGAGCCCTTTTGAAGAAGCAGTTTAAACGCGAGAAGCTGCCAAATGGAATTTGGAGAGA	2061
Qy	2301	GCATCATGTTCAGAGAACAGGTTCACGGGAAGAGCTGGGGAAGTGGGCAGTCACTCTAGCT	2360
Db	2062	GCATCATGTTCAGAGAACAGGTTCACGGGAAGAGCTGGGGAAGTGGGCAGTCACTCTAGCT	2121

QY	2361	TTTCGGGGCAGCATGGAAATCATTTAGGTCTCCTGAGAGAAAGACACCTTGTGACTTCTAT	2420
DB	2122	TTTCGGGCAGCATGGAAATCATTTAGGTCTCCTGAGAGAAAGACACCTTGTGACTTCTAT	2181
QY	2421	AGACAAATTTTCTTTCTTTCTTTTTCACAAAAAATCCCTGTAAATCTGAAATATATATATGT	2480
DB	2182	AGACAAATTTTCTTTCTTTTTCACAAAAAATCCCTGTAAATCTGAAATATATATATGT	2241
QY	2481	ACATACATATATATTTTGGAAAAATGGAGCTATGGTGTAAAGCAACAGGTGGATCAACC	2540
DB	2242	ACATACATATATATTTTGGAAAAATGGAGCTATGGTGTAAAGCAACAGGTGGATCAACC	2301
QY	2541	CAGTTGTACTCTCTTAACATCTGCATTTGAGAGATCAGCTAATACCTTCTCAACAAAA	2600
DB	2302	CAGTTGTACTCTCTTAACATCTGCATTTGAGAGATCAGCTAATACCTTCTCAACAAAA	2361
QY	2601	ATGGAAGGCGAGATGCTAGAAATCCCCCTCAGACGGAGAAAAACATTTTATTCAGTGAAT	2660
DB	2362	ATGGAAGGCGAGATGCTAGAAATCCCCCTCAGACGGAGAAAAACATTTTATTCAGTGAAT	2421
QY	2661	TACACATCTCTTTGTTCTTAAAAAGCAAGTGCTTTGGTGTGGAGACAAAAATCCCTT	2720
DB	2422	TACACATCTCTTTGTTCTTAAAAAGCAAGTGCTTTGGTGTGGAGACAAAAATCCCTT	2481
QY	2721	ACCATTTTCCAGTTGTGCTACTAAGAGATCTCAAAATATTAGTCTTTGTCCGGACCCCTT	2780
DB	2482	ACCATTTT-CACGTTGTGCTACTAAGAGATCTCAAAATATTAGTCTTTGTCCGGACCCCTT	2540
QY	2781	CATAGTACACCTTAGCCCTCAGACTGAGCCAGCTTTGGGGGTCAAGTAGGTAGACCCCTGTT	2840
DB	2541	CATAGTACACCTTAGCCCTCAGACTGAGCCAGCTTTGGGGGTCAAGTAGGTAGACCCCTGTT	2600
QY	2841	AGGACAGAGCGCTAGTGGTAAATCCAAAGAAATGATCCTATCCAAAGCTGATTCAAAA	2900
DB	2601	AGGACAGAGCGCTAGTGGTAAATCCAAAGAAATGATCCTATCCAAAGCTGATTCAAAA	2660
QY	2901	CCCACGCTCACCTTGACAGCCGAGGACACGAGCATCACTCTGCTGACGACCATTAGGG	2960
DB	2661	CCCACGCTCACCTTGACAGCCGAGGACACGAGCATCACTCTGCTGACGACCATTAGGG	2720
QY	2961	GCCTTGCCAAAGGTACTCTTAGAGCAAAACCCAGTACCTCAGACGAGAAAGTCGGGGCTTT	3020
DB	2721	GCCTTGCCAAAGGTACTCTTAGAGCAAAACCCAGTACCTCAGACGAGAAAGTCGGGGCTTT	2780
QY	3021	GACCACATACCATATCTGGTACCCATTTTCTAGGCATTTGTGAATAGGTAGCTAGCTAGTC	3080
DB	2781	GACCACATACCATATCTGGTACCCATTTTCTAGGCATTTGTGAATAGGTAGCTAGCTAGTC	2840
QY	3081	ACACTTTTCAGACCAATTCAAATCTATGCAAAAAATCCCGTGGGCCCTAGATGGAGA	3140
DB	2841	ACACTTTTCAGACCAATTCAAATCTATGCAAAAAATCCCGTGGGCCCTAGATGGAGA	2900
QY	3141	TAATTTTCTTTCTTTCTTTCTTTTATGAAGAGGAAACTGTCTAGGATTCAGACTGAA	3200
DB	2901	TAATTTTCTTTCTTTCTTTTATGAAGAGGAAACTGTCTAGGATTCAGACTGAA	2960
QY	3201	CCACGAGAACCTGGCAACATCAGATTTAAGCTTAAGGTTGGAGGCTAACGAGCTCTACC	3260
DB	2961	CCACGAGAACCTGGCAACATCAGATTTAAGCTTAAGGTTGGAGGCTAACGAGCTCTACC	3020
QY	3261	TCCCTCTTTGTAAATCAAGAAATGTTTAAATGGGATTCGTAATCCTTTAAATAAAGAT	3320
DB	3021	TCCCTCTTTGTAAATCAAGAAATGTTTAAATGGGATTCGTAATCCTTTAAATAAAGAT	3080
QY	3321	GAACCTTGGTTTC 3332	
DB	3081	GAACCTTGGTTTC 3092	
RESULT 15			
AX278461			
LOCUS			
		3059 bp	DNA linear
		AX278461	PAT 02-NOV-2001

RESULT 15	
AX278461	
LOCUS	
DEFINITION	



AX278461  
VERSION AX278461.1 GI:16605915  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
1  
REFERENCE Duecker, K.  
AUTHORS  
TITLE Identification of a dual specificity phosphatase: dusp-10  
JOURNAL Patent: WO 017340-A 1 18-OCT-2001;  
MERCK PATENT-GmbH (DE)  
FEATURES  
source  
1..3059  
/location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
127..2124  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD10545.1"  
/db\_xref="GI:16605915"  
/db\_xref="EMBL:CAD10545"  
/translation="MAHMIQIIVTERLVALLESGETKVLILDSRPFVYNTSHILE  
AININCSKLMKRLQDLVILITELIHSKAKVHIDCSQKVVYVDSQDVLSLSDC  
FLTVLLGLKLEKSFVHLLAGFAEFSPGLCEKSTLVPTCISOPCLPVANIGPT  
RILNLYLGCORFVNLKMGQNGIGVYLNASNTCPKPDFPESHPLRVVNDSPCEK  
ILPLDKSVDFLEKAKSNGVVLHCLAGISRSATIAIATMKMDMLDEAVRFQKE  
KAPTSPNPLFQLLDYEKKIKNOTGASGPKSKILLHLEKNEFPVPAVSEGGKSE  
TPLSPPCADATSEAGQRPVHPASVPSPVQPSLLEDFLVQALSGHLSDRLD  
SNKLRSPSLDTKSVYSASMAASLHGFSSSEDALEYKPSITLDGNKLCQSPVOE  
LSEQTPETSPDKEEASIPKLTQARPSDSQSKRLHVSRTSSGTAQRSLSPHRS  
VEDNHTFPLFLGSLTQOHLTKSAGLKGWHSIDILAPQTSPLTSLTSMYFATESHF  
YSASALYGSASYSAYSCSLPTCGDQVYSVRRQKPSDRADSRHSWHESSPFEKQFK  
RSCQMEFGESIMSENRSEELKGVGSQSSFSGSMELIEVS"  
ORIGIN  
Query Match 81.3%; Score 2710; DB 6; Length 3059;  
Best Local Similarity 94.6%; Pred. No. 0;  
Matches 2895; Conservative 0; Mismatches 0; Indels 165; Gaps 2;  
QY 436 CAGTACACACGAGCTGACCTCATACCTTTAGTACAAATGAGTGGCTGAGCCTTTGAG 495  
DB 1 CAGTACACACGAGCTGACCTCATACCTTTAGTACAAATGAGTGGCTGAGCCTTTGAG 60  
QY 496 CACACACCATTTACATCATCGTGGCAATTAAGAGGAGGTGGAAAGAGGACTTATT 555  
DB 61 CACACACCATTTACATCATCGTGGCAATTAAGAGGAGGTGGAAAGAGGACTTATT 120  
QY 556 GTTGTATGCCCCATGAGATGAGTGAATGGAATGGAATGGAATGGAATGGAATGGAATG 615  
DB 121 GTTGTATGCCCCATGAGATGAGTGAATGGAATGGAATGGAATGGAATGGAATGGAATG 180  
QY 616 CTGGAAGTGGAAACGGAAGTGTCTGCTAAATGATAGCCGCCATTTGTGGAATACAAAT 675  
DB 181 CTGGAAGTGGAAACGGAAGTGTCTGCTAAATGATAGCCGCCATTTGTGGAATACAAAT 240  
QY 676 ACATCCCAATTTTGAAGCAATTAATCAACTGCTCCAGCTTATGAGCGAAGTTG 735  
DB 241 ACATCCCAATTTTGAAGCAATTAATCAACTGCTCCAGCTTATGAGCGAAGTTG 300  
QY 736 CAACAGGACAAAGTGTAAATACAGAGCTCATCCAGCATTCAGCGAAACATAAGGTTGAC 795  
DB 301 CAACAGGACAAAGTGTAAATACAGAGCTCATCCAGCATTCAGCGAAACATAAGGTTGAC 360  
QY 796 ATTGATTGAGTCAGAGTGTGAGTTTACGATCAAAAGCTCCCAAGATGTTGCCTCTCTC 855  
DB 361 ATTGATTGAGTCAGAGTGTGAGTTTACGATCAAAAGCTCCCAAGATGTTGCCTCTCTC 420  
QY 856 TCTTCAGACTGTTTCTCATGCTGACTTCTGGGTAAACTGGAAGAGCTTCAACTCTGTT 915  
DB 421 TCTTCAGACTGTTTCTCATGCTGACTTCTGGGTAAACTGGAAGAGCTTCAACTCTGTT 480

QY 916 CACCTGCTTGC----- 926  
DB 481 CACCTGCTTGCAGGTGGGTTGCTGAGTTCTCTCGTTGTTTCCCTGCGCTCTCTGAGGA 540  
QY 927 ----- 926  
DB 541 AAATCCACTCTAGTCCCTACTGCTATTTCTCAGCTTGTCTTACCTGTTGCAACATTTGG 600  
QY 927 -----AGGAG 931  
DB 601 CCAACCCGAAATTTCTCCCAATTTTATCTTGGTGCAGGAGATGCTCTCAACAGGAG 660  
QY 932 CTGATGCAGCAGAAATGGGATTTGTTATGCTGTTAAATGCCAGCAATACCTGTCCAAAGCCT 991  
DB 661 CTGATGCAGCAGAAATGGGATTTGTTATGCTGTTAAATGCCAGCAATACCTGTCCAAAGCCT 720  
QY 992 GACTTTATCCCGAGTCTCAATTTCTCGGTGCTGCTGTAATGACAGCTTTTGTGAGAA 1051  
DB 721 GACTTTATCCCGAGTCTCAATTTCTCGGTGCTGCTGTAATGACAGCTTTTGTGAGAA 780  
QY 1052 ATTTTGGCGTGTGGCAAAATCAGTAGATTTTCAATGAGAAAGCAAAAGCCTCCAATGGA 1111  
DB 781 ATTTTGGCGTGTGGCAAAATCAGTAGATTTTCAATGAGAAAGCAAAAGCCTCCAATGGA 840  
QY 1112 TGTGTTTGTAGTGCATCTGTTTGTAGTGGATCTCCGCTCCGCCACCATCGCTATCGCTAC 1171  
DB 841 TGTGTTTGTAGTGCATCTGTTTGTAGTGGATCTCCGCTCCGCCACCATCGCTATCGCTAC 900  
QY 1172 ATCTGAGAGGATGGGATGCTCTTTAGTGAAGCTTACAGATTTGTGAAAGAAAAAGA 1231  
DB 901 ATCTGAGAGGATGGGATGCTCTTTAGTGAAGCTTACAGATTTGTGAAAGAAAAAGA 960  
QY 1232 CCTACTATATCTCAAACTTCAATTTTCTGGGCAACTCTGAGCATATGAGAGAGATT 1291  
DB 961 CCTACTATATCTCAAACTTCAATTTTCTGGGCAACTCTGAGCATATGAGAGAGATT 1020  
QY 1292 AGAACACAGATGAGCATCAGGCCCAAGAGCAAACTCAAGCTGCTGCACTGGAGAG 1351  
DB 1021 AGAACACAGATGAGCATCAGGCCCAAGAGCAAACTCAAGCTGCTGCACTGGAGAG 1080  
QY 1352 CCAATTAACCTGCTCTGCTGCTCTCAGAGGTTGAGCAAGAAAGCGAGAGCGCCCTCAGT 1411  
DB 1081 CCAATTAACCTGCTCTGCTGCTCTCAGAGGTTGAGCAAGAAAGCGAGAGCGCCCTCAGT 1140  
QY 1412 CCACCTGTGCGGACTGTCTACCTCAGAGCGAGCAAGAAAGCGCCCTGCTATCCGCC 1471  
DB 1141 CCACCTGTGCGGACTGTCTACCTCAGAGCGAGCAAGAAAGCGCCCTGCTATCCGCC 1200  
QY 1472 AGCGTGCAGCGTGCAGCGTGCAGCGTGCAGCGTGCAGCGTGCAGCGTGCAGCGTGCAG 1531  
DB 1201 AGCGTGCAGCGTGCAGCGTGCAGCGTGCAGCGTGCAGCGTGCAGCGTGCAGCGTGCAG 1260  
QY 1532 GGGCTCAGTGGGCTGACCTGTCCGAGACAGCGTGCAGCGTGCAGCGTGCAGCGTGCAGCGT 1591  
DB 1261 GGGCTCAGTGGGCTGACCTGTCCGAGACAGCGTGCAGCGTGCAGCGTGCAGCGTGCAGCGT 1320  
QY 1592 TCCTTCTCTCTGAGATCAAAATCAGTTTCAATTCAGCGAGCATGGCAGCATCCTTACAT 1651  
DB 1321 TCCTTCTCTCTGAGATCAAAATCAGTTTCAATTCAGCGAGCATGGCAGCATCCTTACAT 1380  
QY 1652 GGCCTTCTCTCATCAGAGATGCTTTGGAATPACAAACCTTCCACTACTCTGGATGGG 1711  
DB 1381 GGCCTTCTCTCATCAGAGATGCTTTGGAATPACAAACCTTCCACTACTCTGGATGGG 1440  
QY 1712 ACCAACAGCTATGCGCTTCTCCCTGTTTCCAGAACTATTCGAGAGCATGCTCCCGAAACC 1771  
DB 1441 ACCAACAGCTATGCGCTTCTCCCTGTTTCCAGAACTATTCGAGAGCATGCTCCCGAAACC 1500  
QY 1772 AGTCTGTGTAAGAGGAGGAGGAGCATCCCCAAGAGTGTGAGAGCGCGAGGCTTTCAGAC 1831  
DB 1501 AGTCTGTGTAAGAGGAGGAGGAGCATCCCCAAGAGTGTGAGAGCGCGAGGCTTTCAGAC 1560  
QY 1832 AGCCAGAGCAAGCGATTGCAATTCGGTCAGAACACAGCAGCAGTGGCAGCGCCAGGCTCC 1891



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 27, 2004, 18:38:58 ; Search time 1216 Seconds

(without alignments)  
11640.622 Million cell updates/sec

Title: US-09-964-277-20

Perfect score: 3332

Sequence: 1 gaggaagaggagagataata.....ataaagatgaacttggtttc 3332

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_29Jan04:\*

1: Geneseq1980s:\*

2: Geneseq1990s:\*

3: Geneseq2000s:\*

4: Geneseq2001as:\*

5: Geneseq2001bs:\*

6: Geneseq2002s:\*

7: Geneseq2003as:\*

8: Geneseq2003bs:\*

9: Geneseq2003cs:\*

10: Geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3332	100.0	3332	6	ABK48378
2	3158	94.8	3496	6	ABK47596
3	3118.8	93.6	3766	6	ABK14474
4	3106.2	93.2	5450	6	ACC60559
5	3106.2	93.2	5450	6	ACC60572
6	2950	88.5	3544	5	RAA14639
7	2950	88.5	3544	6	ABK49402
8	2950	88.5	5145	5	ABV20833
9	2950	88.5	5145	5	ABV21080
10	2950	88.5	5145	5	ABV21080
11	2950	88.5	5145	5	ABV20978
12	2950	88.5	5145	5	ABV21092
13	2950	88.5	5145	5	ABV21312
14	2950	88.5	5145	5	ABV21316
15	2950	88.5	5145	5	ABV26826
16	2950	88.5	5145	5	ABV27131
17	2950	88.5	5145	5	ABV26923
18	2950	88.5	5145	5	ABV27135
19	2950	88.5	5145	5	ABV28657
20	2950	88.5	5145	5	ABV28657
21	2950	88.5	5145	5	ABV26934
22	2924.6	87.8	5111	6	ACC60521
23	2766.4	83.0	4790	6	ABN83966

24	2717.2	81.5	3104	6	ABN59704	Novel hum
25	2710	81.3	3059	6	AAAS15768	Novel hum
26	2602.8	78.1	2966	4	AAH99685	Human pro
27	2390.8	71.8	2732	4	AAD09492	Human SGP
28	1762.4	52.9	2102	7	ADA53105	Human cod
29	1741	52.3	2118	4	AAF30479	Human pro
30	1638.2	49.2	2756	6	ACC60560	Polynucle
31	886.8	26.6	2558	9	ADB33596	Human CDN
32	595	17.9	787	4	AAI95265	Human neu
33	595	17.9	787	4	AAI98142	Human neu
34	537.4	16.1	877	4	AAAS40844	CDNA enco
35	537.4	16.1	877	4	ABA06492	Human CDN
36	537.4	16.1	877	6	ABV83829	Human pol
37	431.4	12.9	515	4	AAH11589	Human CDN
38	417.2	12.5	425	5	ABV10726	Human pro
39	415.2	12.5	467	5	ABV40998	Human pro
40	415.2	12.5	467	5	ABV31891	Human pro
41	415.2	12.5	467	5	ABV40849	Human pro
42	409.2	12.3	478	8	ACH38589	Human end
43	408	12.2	408	6	ACC60561	Polynucle
44	401.2	12.0	408	5	ABV10907	Human pro
45	399.2	12.0	438	5	ABV32059	Human pro

## ALIGNMENTS

RESULT 1

ABK48378

ID ABK48378 standard; cDNA; 3332 BP.

XX

AC ABK48378;

XX

DT 02-JUL-2002 (first entry)

XX

DE cDNA encoding human DSP-16 alternative form protein.

XX

KW Human; dual-specificity phosphatase-3; DSP-16; autoimmune disease;  
KW mitogen-activated protein kinase; MAP; Duchenne muscular dystrophy;  
KW cancer; graft-versus-host disease; allergy; metabolic disease;  
KW abnormal cell growth; abnormal cell proliferation; contact inhibition;  
KW cell cycle abnormality; anchorage independent cell growth; apoptosis;  
KW intercellular adhesion; DSP-16 modulator; chromosome 12p; gene; ss.

XX

OS Homo sapiens.

XX

PH Key Location/Qualifiers

FT CDS

FT 842..2395

FT /\*tag= a

FT /product= "Human dual-specificity phosphatase-3 (DSP-16)

FT alternative form protein"

FT WO200226997-A2.

XX

PD 04-APR-2002.

XX

PF 25-SEP-2001; 2001WO-US030124.

XX

PR 26-SEP-2000; 2000US-0235487P.

XX

PA (CEPT-) CEPTVR INC.

XX

PI Luche RM, Wei B;

XX

DR WPI; 2002-315802/35.

XX

P-PSDB; AAU79159.

XX

PT New DSP-16 polypeptide, useful for identifying modulators of its

PT activity, which can be used in the treatment of disorders such as

PT Duchenne muscular dystrophy, or cancer.

XX

PS Claim 56; Fig 3; 87pp; English.

XX

The present invention relates to a new polypeptide, DSP-16, having a 665 amino acid sequence, given in the specification, or a variant having at least 50 % identical residues, which retains the ability to dephosphorylate an activated mitogen-activated protein (MAP) kinase. The invention can be used for identifying agents which modulate DSP-16 activity, for modulation of a proliferative response in a cell, survival of a cell, or differentiation of a cell. The cell displays contact inhibition of cell growth or anchorage independent growth and may display altered inter-cellular adhesion. The agent may modulate apoptosis, or the cell cycle. The identified modulators can be used to treat Duchenne muscular dystrophy, cancer, graft-versus-host disease, autoimmune diseases, allergies, metabolic diseases, abnormal cell growth, abnormal cell proliferation, and cell cycle abnormalities. The present nucleic acid sequence is that of the human dual-specificity phosphatase-3 (DSP-16), gene located on chromosome 12p. This sequence encodes the human DSP-16 alternative form protein of the invention.

SQ Sequence 3332 BP; 909 A; 805 C; 823 G; 795 T; 0 U; 0 Other;

```
Query Match      100.0%; Score 3332; DB 6; Length 3332;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3332; Conservative 0; Mismatches 0; Indels 0;
```

```

Best local similarity 100.00%, Freq. NO: 0;
Matches 3332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 GAGAGAAGGAGAGATAATATACTGAAAGAGAGGAGAGACCGGACCGGAC 60

Db  
1 GAGAGAAGGAGAAGATAATATCTGAAAAGAAAGAGGAGGAGAGCGGACGGGAC 60

Qy 61 GCGAGCGGAGCGCAGCCGCCCTCTCGGCTCCGCGCGCGCCCTCGCAAGTCCGGAGGC 120

Db 61 GCGAGCGGGAGCGCAGCCGCCCTCTCGGCTCCGCGGCGCGCCTCGCAAGTCCGGAGGC 120

QY 121 GAGGGGGCCGAGGGGAGCGCCGTGACAACCTTCGTTTCCCTCTGAGGGAATTGGGAG 180

D<sub>b</sub> 121 GAGGGGGCCGAGGGGAGACGCCGTGACAACTTTCGTTCCTCTGAGGGAA TTGGGAG 180

QV 181 GTCGGCGGCCCAAAAGCTTTCAGTCCAGTGTAAGCTGTTGGAGCGCGGAGCAAAGGT 240

D<sub>b</sub> 181 GTCGGCGGCCCAAGCTTTCAGTCCAGTGTAAGCTGTTGGAGCGCGGAGCAAGGT 240

241 AAAGAAATGATGTAATGCCGCTGGCTGCTCAAAGCACTCTTTGTGTGGGAATGGTTATTCC 300

bp 241 AAAGAATGATGTAATGCCGCTGGCGTCCAAAGCATCTTTTGTGTGGAA TGGTTATTCC 300

OV 301 AGTCATCTCTTTATGCAATCAAA TGTGAGGGGCTGCTTTGTGGA CGGAGTCTTTTGCAAGA 360

301 AGTCATCTCTTTATGAATCAAAATGTGAGGGGCTGCTTTGTGGA CGGAGTCCTTTGCAAGA 360

361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000

[illegible][illegible]

421  
480

[illegible]

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

100

.....

100

Downloaded from <http://ajphaphapublications.org/> on 01/01/2017

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

.....

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

721	DB	ATGAAGCGAAGGTTGCAACAGGACAAAGTTGTTAATTA	CAGAGCTCATCCAGCATTCACGG	7260
781	QY	AAACATAAGGTTGACATTGATTCGAGTCAGAAAGTGTAGTTT	TACGATCAAAGCTCCCAA	840
781	DB	AAACATAAGGTTGACATTGATTCGAGTCAGAAAGTGTAGTTT	TACGATCAAAGCTCCCAA	840
841	QY	GATGTTGCCTCTCTCTCTTCAGACTGTTTTCTCACTGTACTTCT	TGGGTAAACTCGAGAAG	900
841	DB	GATGTTGCCTCTCTCTCTTCAGACTGTTTTCTCACTGTACTTCT	TGGGTAAACTCGAGAAG	900
901	QY	AGCTTTCAACTCTGTTCACTCTGTTGCAGGAGCTGATCAGCAGAA	TGGGATTTGTTATGT	960
901	DB	AGCTTTCAACTCTGTTCACTCTGTTGCAGGAGCTGATCAGCAGAA	TGGGATTTGTTATGT	960
961	QY	GTTAAATGCCAGCAATACCTGTCCAAAGCCTGACTTTATCCC	GAGTCTCATTTCTCGCG	1020
961	DB	GTTAAATGCCAGCAATACCTGTCCAAAGCCTGACTTTATCCC	GAGTCTCATTTCTCGCG	1020
1021	QY	TGTCGCTGTCAATGACAGCTTTTGTGAGAAAATTTTGCCTGCT	TGGACAAATCAGTAGA	1080
1021	DB	TGTCGCTGTCAATGACAGCTTTTGTGAGAAAATTTTGCCTGCT	TGGACAAATCAGTAGA	1080
1081	QY	TTTCAATTGAGAAAGCAAAAGCCTCCAAATGATGTGTTCTAGT	GCACTGTTTGTAGCTGGGAT	1140
1081	DB	TTTCAATTGAGAAAGCAAAAGCCTCCAAATGATGTGTTCTAGT	GCACTGTTTGTAGCTGGGAT	1140
1141	QY	CTCCGCTCCGCGACCAACCTCGCTATCGCTACATCATGAAGAG	GATGACATGCTCTTTAGA	1200
1141	DB	CTCCGCTCCGCGACCAACCTCGCTATCGCTACATCATGAAGAG	GATGACATGCTCTTTAGA	1200
1201	QY	TGAAGCTTTACAGATTTGTGAAAGAAAAGACCTACTATATCT	CCAAACTTCATTTTCT	1260
1201	DB	TGAAGCTTTACAGATTTGTGAAAGAAAAGACCTACTATATCT	CCAAACTTCATTTTCT	1260
1261	QY	GGGCGCAACTCCTCGACTATGAGNAGAGATTAAAGAACCCAG	ACTGGAGCATCAGGGCCAAA	1320
1261	DB	GGGCGCAACTCCTCGACTATGAGNAGAGATTAAAGAACCCAG	ACTGGAGCATCAGGGCCAAA	1320
1321	QY	GAGCAAAACTCAAGCTGCTGCACTCGGAGAAAGCCAAATGAA	CTGTCCTGCTGCTCTCAGA	1380
1321	DB	GAGCAAAACTCAAGCTGCTGCACTCGGAGAAAGCCAAATGAA	CTGTCCTGCTGCTCTCAGA	1380
1381	QY	GGGTGGACAGAAAAGCGAGAGCGCCCTCAGTCCACCCTGTG	CGCACTCTGCTACTCTCAGA	1440
1381	DB	GGGTGGACAGAAAAGCGAGAGCGCCCTCAGTCCACCCTGTG	CGCACTCTGCTACTCTCAGA	1440
1441	QY	GGCAGCAGGACAAAGCCCGTGCACTCCGCGCAGCGTGCCAG	CGTGCCACGCGTGCAGCC	1500
1441	DB	GGCAGCAGGACAAAGCCCGTGCACTCCGCGCAGCGTGCCAG	CGTGCCACGCGTGCAGCC	1500
1501	QY	CTCGCTCTTTAGAGGACAGCGCCGCTGGTACAGCGCTCAGT	GGGTGCACTGTCCGCGAGA	1560
1501	DB	CTCGCTCTTTAGAGGACAGCGCCGCTGGTACAGCGCTCAGT	GGGTGCACTGTCCGCGAGA	1560
1561	QY	CAGGCTGGAGAGACGCAATAAGCTCAAGGGTTCCTTCTCTG	TGGATATCAAAATCAGTTTC	1620
1561	DB	CAGGCTGGAGAGACGCAATAAGCTCAAGGGTTCCTTCTCTG	TGGATATCAAAATCAGTTTC	1620
1621	QY	ATATTACGCGCATGGCAGCATCTTTACATGGCTTCTCTCAT	CATCAGAAAGATGCTTTTGA	1680
1621	DB	ATATTACGCGCATGGCAGCATCTTTACATGGCTTCTCTCAT	CATCAGAAAGATGCTTTTGA	1680
1681	QY	ATACTACAAACCTTCCACTACTCTGGAATGGNACCAACAG	CTATGCGCAAGTCTCTCCCTGT	1740
1681	DB	ATACTACAAACCTTCCACTACTCTGGAATGGNACCAACAG	CTATGCGCAAGTCTCTCCCTGT	1740
1741	QY	TCAGGAACCTATCGGACAGACTCCCGGAAACCAAGCTCTG	ATAAGAGGAAGCCAGCATCCC	1800
1741	DB	TCAGGAACCTATCGGACAGACTCCCGGAAACCAAGCTCTG	ATAAGAGGAAGCCAGCATCCC	1800
1801	QY	CAAGAAGCTCAGACCGCGACGCTTTCAAGCAGCAGCAAG	CGGATTCGATTCGCTCAG	1860
1801	DB	CAAGAAGCTCAGACCGCGACGCTTTCAAGCAGCAGCAAG	CGGATTCGATTCGCTCAG	1860

1861 AACAGCAGCAGTGGSCACCGCCAGAGTCCCTTTTATCTCCACTGCATCGAAGTGGAG 1920  
Db  
1861 AACAGCAGCAGTGGSCACCGCCAGAGTCCCTTTTATCTCCACTGCATCGAAGTGGAG 1920  
Qy 1921 CGTGAGGACAAATTACCAACACCAAGCTTCCTTTTCGGCCCTTTCCACAGCAGCAGCACCT 1980  
Db 1921 CGTGAGGACAAATTACCAACACCAAGCTTCCTTTTCGGCCCTTTCCACAGCAGCAGCACCT 1980  
Qy 1981 CACGAAGTCTGCTGGCCGTGGGCTTTAGGGCTGCGACTCGGATATCTTGSCCCCCAGAC 2040  
Db 1981 CACGAAGTCTGCTGGCCGTGGGCTTTAGGGCTGCGACTCGGATATCTTGSCCCCCAGAC 2040  
Qy 2041 CTCTACCCCTTCCCTGACACAGCAGCTGGTATTTTGGCCACAGAGTCTCTCACTTCTACTC 2100  
Db 2041 CTCTACCCCTTCCCTGACACAGCAGCTGGTATTTTGGCCACAGAGTCTCTCACTTCTACTC 2100  
Qy 2101 TGCCTCAGCCATCTACGGAGCAGTGCAGTACTCTGCTCTAGCTGAGCAGCAGTGC 2160  
Db 2101 TGCCTCAGCCATCTACGGAGCAGTGCAGTACTCTGCTCTAGCTGAGCAGCAGTGC 2160  
Qy 2161 CACTTGCAGGACCAAGTCTATTCTGTGGCAGCGCAGCAAGCAAGTGACAGAGCTGA 2220  
Db 2161 CACTTGCAGGACCAAGTCTATTCTGTGGCAGCGCAGCAAGCAAGTGACAGAGCTGA 2220  
Qy 2221 CTCGGCGGAGCTGGCATGAAGAGACCCCTTTGAAAGCAGCTTTAAACGCAAGCTG 2280  
Db 2221 CTCGGCGGAGCTGGCATGAAGAGACCCCTTTGAAAGCAGCTTTAAACGCAAGCTG 2280  
Qy 2281 CCAATGGAATTTGGAGAGAGCATGTCTAGAGACAGCTGAAATCAITGAGTCTCTCGAGA 2340  
Db 2281 CCAATGGAATTTGGAGAGAGCATGTCTAGAGACAGCTGAAATCAITGAGTCTCTCGAGA 2340  
Qy 2341 AGTGGGAGTCACTGCTAGCTTTTGGGAGCAGTGGAAATCAITGAGTCTCTCGAGA 2400  
Db 2341 AGTGGGAGTCACTGCTAGCTTTTGGGAGCAGTGGAAATCAITGAGTCTCTCGAGA 2400  
Qy 2401 AAGACACTGTGACTCTATAGACAATTTTTTTCTGTGTCTCAAAAAAATTCCTGTGA 2460  
Db 2401 AAGACACTGTGACTCTATAGACAATTTTTTTCTGTGTCTCAAAAAAATTCCTGTGA 2460  
Qy 2461 AATCTGAAATATATATATGTACATATATATTTTTTGGMAATGGAGCTATGGTGA 2520  
Db 2461 AATCTGAAATATATATATGTACATATATATTTTTTGGMAATGGAGCTATGGTGA 2520  
Qy 2521 AAGCAACAGGTGGATCAACCCAGTGTACTCTCTTAAACATCTGCATTTGAGAGATCAGC 2580  
Db 2521 AAGCAACAGGTGGATCAACCCAGTGTACTCTCTTAAACATCTGCATTTGAGAGATCAGC 2580  
Qy 2581 TAATACTTCTCTCAACAAATGMAAGGCGAGATGCTAGAAATCCCTCCCTAGACGGAGAA 2640  
Db 2581 TAATACTTCTCTCAACAAATGMAAGGCGAGATGCTAGAAATCCCTCCCTAGACGGAGAA 2640  
Qy 2641 AACCATTTTATTCAGTGAATTAACATCTCTTGTCTTTAAAAAGCAAGTGTCTTTGGT 2700  
Db 2641 AACCATTTTATTCAGTGAATTAACATCTCTTGTCTTTAAAAAGCAAGTGTCTTTGGT 2700  
Qy 2701 GTTGGAGACAAATCCCTTACCATTTTCCAGTGTGTCTCTAAGAGATCTCAAAATTT 2760  
Db 2701 GTTGGAGACAAATCCCTTACCATTTTCCAGTGTGTCTCTAAGAGATCTCAAAATTT 2760  
Qy 2761 AGTCTTTGTCTCGGACCCCTTCCATAGTACACTTTAGCGCTGAGACTGAGCAGCTTGGGG 2820  
Db 2761 AGTCTTTGTCTCGGACCCCTTCCATAGTACACTTTAGCGCTGAGACTGAGCAGCTTGGGG 2820  
Qy 2821 TCAGGTAGGTAGACCTTGTAGGAGCAGAGCTTGTGGTAAATCCAGAGAAATGATCCT 2880  
Db 2821 TCAGGTAGGTAGACCTTGTAGGAGCAGAGCTTGTGGTAAATCCAGAGAAATGATCCT 2880  
Qy 2881 ATCCAAAGCTGATTACAAACCCACCGTCACTGACAGCGGAGGACACGACATCACTC 2940  
Db 2881 ATCCAAAGCTGATTACAAACCCACCGTCACTGACAGCGGAGGACACGACATCACTC 2940

2941 TGCTGGACGGACCAATTAGGGCTTGGCAGGCTTACCTTAGAGCAAAACCCAGTACTCA 3000  
Db 2941 TGCTGGACGGACCAATTAGGGCTTGGCAGGCTTACCTTAGAGCAAAACCCAGTACTCA 3000  
Qy 3001 GACAGGAAAGTCGGGGCTTTTGACCACTTACCATATCTGGTAGCCCATTTCTAGGCATTGT 3060  
Db 3001 GACAGGAAAGTCGGGGCTTTTGACCACTTACCATATCTGGTAGCCCATTTCTAGGCATTGT 3060  
Qy 3061 GAATAGTAGGTAGTACACTTTTTCAGACCAATTCMACTGTCTATGCAAAAAAT 3120  
Db 3061 GAATAGTAGGTAGTACACTTTTTCAGACCAATTCMACTGTCTATGCAAAAAAT 3120  
Qy 3121 CCCGTGGCCCTAGATGAGATAAATTTTTTTCTCTCAGCTTTATGAAGAGAGGAAA 3180  
Db 3121 CCCGTGGCCCTAGATGAGATAAATTTTTTTCTCTCAGCTTTATGAAGAGAGGAAA 3180  
Qy 3181 CTGTCTAGGATTCAGCTGAAACCAAGGAACTGGGCAACATCAGATTTAAGCTAAGGTT 3240  
Db 3181 CTGTCTAGGATTCAGCTGAAACCAAGGAACTGGGCAACATCAGATTTAAGCTAAGGTT 3240  
Qy 3241 GGGAGGCTTAACGAGTCTTACCTCCCTCTTTGTTAAATCAAGAAATTTGTTAAATGGGATTG 3300  
Db 3241 GGGAGGCTTAACGAGTCTTACCTCCCTCTTTGTTAAATCAAGAAATTTGTTAAATGGGATTG 3300  
Qy 3301 TCAATCCTTTTAAATAAAGATGAACCTTGGTTTC 3332  
Db 3301 TCAATCCTTTTAAATAAAGATGAACCTTGGTTTC 3332  
RESULT 2  
ABK47596  
ID ABK47596 standard; cDNA; 3496 BP.  
XX  
AC ABK47596;  
XX  
DT 02-JUL-2002 (first entry)  
XX  
DE cDNA encoding human dual-specificity phosphatase-3 (DSP-16) protein.  
XX  
KW Human; dual-specificity phosphatase-3; DSP-16; autoimmune disease;  
KW mitogen-activated protein kinase; MAP; Duchenne muscular dystrophy;  
KW cancer; graft-versus-host disease; allergy; metabolic disease;  
KW abnormal cell growth; abnormal cell proliferation; contact inhibition;  
KW cell cycle abnormality; anchorage independent cell growth; apoptosis;  
KW intercellular adhesion; DSP-16 modulator; chromosome 12p; gene; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 582..2559  
FT /\*tag= a  
FT /product= "Human dual-specificity phosphatase-3 (DSP-16)  
FT protein"  
XX  
XX WO200226997-A2.  
XX  
XX 04-APR-2002.  
XX  
XX 25-SEP-2001; 2001WO-US030124.  
XX  
XX 26-SEP-2000; 2000US-0235487P.  
XX  
XX (CEPT-) CEPTYR INC.  
XX  
XX Luche RM, Wei B;  
XX WPI; 2002-315802/35.  
XX P-PSDB; AAU79156.  
XX  
XX New DSP-16 polypeptide, useful for identifying modulators of its  
XX activity, which can be used in the treatment of disorders such as  
XX Duchenne muscular dystrophy, or cancer.  
XX



Db 1801 GCAGCATCCTTACATGGCTTCTCCTCATCAGAAGATGCTTTGGAAATACTACAAGACCTTCC 1860  
QY 1697 ACTACTCTGGATGGACCAACAAAGCTATGCCAGTTCTCCCTGTTCAGGAACATATCGGAG 1756  
Db 1861 ACTACTCTGGATGGACCAACAAAGCTATGCCAGTTCTCCCTGTTCAGGAACATATCGGAG 1920  
QY 1757 CAGACTCCCGAAACAGTCCCTGATTAAGAGGAGAGCCAGCATCCCCAGAGAGCTGCAGACC 1816  
Db 1921 CAGACTCCCGAAACAGTCCCTGATTAAGAGGAGAGCCAGCATCCCCAGAGAGCTGCAGACC 1980  
QY 1817 GCCAGGCTTCAGACAGCCAGAGCAAGCGATTGCAATTCGGTTCAGAACCCAGCAGTGGC 1876  
Db 1981 GCCAGGCTTCAGACAGCCAGAGCAAGCGATTGCAATTCGGTTCAGAACCCAGCAGTGGC 2040  
QY 1877 ACCGCCACAGAGTCCCTTTATCTCCATCTCATCGAAGTGGAGCGTGGAGGACAATTAC 1936  
Db 2041 ACCGCCACAGAGTCCCTTTATCTCCATCTCATCGAAGTGGAGCGTGGAGGACAATTAC 2100  
QY 1937 CACACAGCTTCTTTTCGGCCCTTTCCACCAGCAGCAGCACCTCAGCAAGTCTGCTGGC 1996  
Db 2101 CACACAGCTTCTTTTCGGCCCTTTCCACCAGCAGCAGCACCTCAGCAAGTCTGCTGGC 2160  
QY 1997 CTGGGCTTAAGGCTGGCACTCGGATATCTTGGCCGCCAGACCTCTACCCCTTCCCTG 2056  
Db 2161 CTGGGCTTAAGGCTGGCACTCGGATATCTTGGCCGCCAGACCTCTACCCCTTCCCTG 2220  
QY 2057 ACCACAGCTGTATTTTGGCCACAGAGTCCCTCACACTTCTACTCTGCTCAGCCACTAC 2116  
Db 2221 ACCACAGCTGTATTTTGGCCACAGAGTCCCTCACACTTCTACTCTGCTCAGCCACTAC 2280  
QY 2117 GGAGCAGTGGCAGTTACTTGCCTACAGTGCAGCCAGCTGCCACTTGCAGGAGCCAA 2176  
Db 2281 GGAGCAGTGGCAGTTACTTGCCTACAGTGCAGCCAGCTGCCACTTGCAGGAGCCAA 2340  
QY 2177 GTCATTTCTGCGCAGCGGCAGAGCCAGTGCAGAGTCACTGCTCGCGCGAGCTGG 2236  
Db 2341 GTCATTTCTGCGCAGCGGCAGAGCCAGTGCAGAGTCACTGCTCGCGCGAGCTGG 2400  
QY 2237 CATGAAGAGAGCCCTTTGAAAAGCAGTTTAAACGCAAGAGCTGCCAAATGGAAATTTGA 2296  
Db 2401 CATGAAGAGAGCCCTTTGAAAAGCAGTTTAAACGCAAGAGCTGCCAAATGGAAATTTGA 2460  
QY 2297 GAGACATCATGTACAGAAACAGTACGCGGAAGAGCTGGGAAAGTGGGCACTCAGTCT 2356  
Db 2461 GAGACATCATGTACAGAAACAGTACGCGGAAGAGCTGGGAAAGTGGGCACTCAGTCT 2520  
QY 2357 AGCTTTTCGGCAGCATGGAATCATTTAGGTCTCCTGAGAAGAAACACACTTGTGACTT 2416  
Db 2521 AGCTTTTCGGCAGCATGGAATCATTTAGGTCTCCTGAGAAGAAACACACTTGTGACTT 2580  
QY 2417 CTATAGACAAATTTTCTTTCTGTTCAAAAAAATTCCTGTAATCTGAAATATATAT 2476  
Db 2581 CTATAGACAAATTTTCTTTCTGTTCAAAAAAATTCCTGTAATCTGAAATATATAT 2640  
QY 2477 ATGTACATACATATATATTTTGGAAATGGAGTATGGTGAAGCAACAGCTGGATC 2536  
Db 2641 ATGTACATACATATATATTTTGGAAATGGAGTATGGTGAAGCAACAGCTGGATC 2700  
QY 2537 AACCCAGTTGTTACTCTCTTAACATCTGCAATTTGAGAGATCAGCTAATCTTCTCAAC 2596  
Db 2701 AACCCAGTTGTTACTCTCTTAACATCTGCAATTTGAGAGATCAGCTAATCTTCTCAAC 2760  
QY 2597 AAAAATGAAGGCGAGATGTAGAAATCCCCCTAGACGGAGGAAACCAATTTATTCAGT 2656  
Db 2761 AAAAATGAAGGCGAGATGTAGAAATCCCCCTAGACGGAGGAAACCAATTTATTCAGT 2820  
QY 2657 GAAATTAACATCTCTTGTCTTAAAGAACAGTGTCTTTGGTGGTGGAGGACAAATC 2716  
Db 2821 GAAATTAACATCTCTTGTCTTAAAGAACAGTGTCTTTGGTGGTGGAGGACAAATC 2880  
QY 2717 CCCTACCAATTTTCCAGCTGTGCTACTAAGAGATCTCAAAATATAGTCTTTGCGGAGCC 2776

Db 2881 CCCTACCAATTTTCCAGCTGTGCTACTAAGAGATCTCAAAATATTAGTCTTTGTCGGAGCC 2940  
QY 2777 CTTCCATAGTACACCTTTAGCGCTGAGACTGAGCCAGCTTGGGGGTCAAGTAGGTAGACCC 2836  
Db 2941 CTTCCATAGTACACCTTTAGCGCTGAGACTGAGCCAGCTTGGGGGTCAAGTAGGTAGACCC 3000  
QY 2837 TGTTAGGACAGAGCGCTAGTGGTAAATCCAGAGAAATGATCTTATCCAAAGCTGATTCA 2896  
Db 3001 TGTTAGGACAGAGCGCTAGTGGTAAATCCAGAGAAATGATCTTATCCAAAGCTGATTCA 3060  
QY 2897 CAAACCCACGCTCACCTGACAGCCGAGGAGCAGAGCATCACTCTGCTGACGAGCCATT 2956  
Db 3061 CAAACCCACGCTCACCTGACAGCCGAGGAGCAGAGCATCACTCTGCTGACGAGCCATT 3120  
QY 2957 AGGGCCCTTGCAGAGCTTACCTTAGAGCAAAACCCAGTACCTCAGACAGGAAAGTCGGG 3016  
Db 3121 AGGGCCCTTGCAGAGCTTACCTTAGAGCAAAACCCAGTACCTCAGACAGGAAAGTCGGG 3180  
QY 3017 CTTTGACCACTTACCATATCTGGTAGCCCATTTCTAGGCATTTCTAGGATTTAGTAGGTAGCT 3076  
Db 3181 CTTTGACCACTTACCATATCTGGTAGCCCATTTCTAGGCATTTCTAGGATTTAGTAGGTAGCT 3240  
QY 3077 AGTCACACTTTTTCAGACCAATTCAACTGTCTATGCAAAAAATTCGCTGGGCTAGATG 3136  
Db 3241 AGTCACACTTTTTCAGACCAATTCAACTGTCTATGCAAAAAATTCGCTGGGCTAGATG 3300  
QY 3137 GAGATAATTTTTTTTTTCTCTCAGCTTTATGAAGAGGGAACCTGCTAGGATTCAGC 3196  
Db 3301 GAGATAATTTTTTTTTTCTCTCAGCTTTATGAAGAGGGAACCTGCTAGGATTCAGC 3360  
QY 3197 TGAACCAACAGGAAACCTGGCAACATCAAGATTTAAGCTAAGGTGGAGGCTAACGAGTC 3256  
Db 3361 TGAACCAACAGGAAACCTGGCAACATCAAGATTTAAGCTAAGGTGGAGGCTAACGAGTC 3420  
QY 3257 TACCTCCCTCTTTGTAATCAAGAAATTTTAAATGGGATTTGCAATCTTTAAATAA 3316  
Db 3421 TACCTCCCTCTTTGTAATCAAGAAATTTTAAATGGGATTTGCAATCTTTAAATAA 3480  
QY 3317 AGATGAACCTTGGTTTC 3332  
Db 3481 AGATGAACCTTGGTTTC 3496  
RESULT 3  
ABK14474  
ID ABK14474 standard; cDNA; 3766 BP.  
XX AC ABK14474;  
XX AC ABK14474;  
DT 08-MAY-2002 (first entry)  
XX Human protein phosphatase 7 (PP7) cDNA sequence.  
DE Human; protein phosphatase 7; PP7; immune system disorder; AIDS;  
KW acquired immunodeficiency syndrome; thymic hypoplasia; anaemia; asthma;  
KW Crohn's disease; neurological disorder; epilepsy; Huntington's disease;  
KW dementia; Parkinson's disease; developmental disorder; Down's syndrome;  
KW cell proliferative disorder; cancer; adenocarcinoma; leukaemia; lymphoma;  
KW melanoma; myeloma sarcoma; gene; ss.  
XX Homo sapiens.  
OS  
XX  
XX  
FH Key Location/Qualifiers  
CDS 538..2535  
FT /\*tag= a  
FT /product= "Protein\_phosphatase\_7 (PP7)"  
XX WO200210363-A2.  
PN  
XX  
XX PD 07-FEB-2002.  
XX 26-JUL-2001; 2001WO-US023716.  
XX PF  
XX







Db 1501 CTTGTCCTCTGCTCTCAGAGGGTGGACAGAAAAGCGAGAGCGCCCTTCAGTCCACCCCTGT 1560  
QY 1421 GCCGACTCTGCTACTCCTCAGAGCGCAGCAGGACAAAGGCCCGTGTGATCCCGCCAGCGTGCC 1480  
Db 1561 GCCGACTCTGCTACTCCTCAGAGCGCAGCAGGACAAAGGCCCGTGTGATCCCGCCAGCGTGCC 1620  
QY 1481 AGCGTGCCCGCGCGCAGCGCTGCTCTGTATAGGACAGCGCGGTGTGTACAGCGCTCAGT 1540  
Db 1621 AGCGTGCCCGCGCGCAGCGCTGCTCTGTATAGGACAGCGCGGTGTGTACAGCGCTCAGT 1680  
QY 1541 GGGTGCACTCTGTCGCGCAGACAGCGCTGSAAGACAGCAATAAGCTCAAGCGTTCCTCTCT 1600  
Db 1681 GGGTGCACTCTGTCGCGCAGACAGCGCTGSAAGACAGCAATAAGCTCAAGCGTTCCTCTCT 1740  
QY 1601 CTGGATATCAATCAGTTTCATATTCAGCCAGCATGCGAGCAGCTTACATGCTTCTCC 1660  
Db 1741 CTGGATATCAATCAGTTTCATATTCAGCCAGCATGCGAGCAGCTTACATGCTTCTCC 1800  
QY 1661 TCATCAGAAGATGCTTTGGAAATCTACAAACCTTCCACTACTCTGGATGGGACCAACAAG 1720  
Db 1801 TCATCAGAAGATGCTTTGGAAATCTACAAACCTTCCACTACTCTGGATGGGACCAACAAG 1860  
QY 1721 CTATGCCAGTTCTCCCTGTTTCAGAACTATCGAGCAGACCTCCCGAAACCAAGCTCTGAT 1780  
Db 1861 CTATGCCAGTTCTCCCTGTTTCAGAACTATCGAGCAGACCTCCCGAAACCAAGCTCTGAT 1920  
QY 1781 AAGGAGGAAGCCAGCATCCCAAGAGCTGCAGACCGCCAGGCTTCAGACAGCCAGAGC 1840  
Db 1921 AAGGAGGAAGCCAGCATCCCAAGAGCTGCAGACCGCCAGGCTTCAGACAGCCAGAGC 1980  
QY 1841 AAGCGATTGCAATTCGGTCAGAACAGCAGCAGTGCGCACCGCCAGAGGTCCTTTTATCT 1900  
Db 1981 AAGCGATTGCAATTCGGTCAGAACAGCAGCAGTGCGCACCGCCAGAGGTCCTTTTATCT 2040  
QY 1901 CCATCTGATCCAGTGGAGCGTGGAGGACAAATTCACACAGCAGCTTCTTTTCGCGCTT 1960  
Db 2041 CCATCTGATCCAGTGGAGCGTGGAGGACAAATTCACACAGCAGCTTCTTTTCGCGCTT 2100  
QY 1961 TCCACAGCCAGCAGCAGCTCAGAAAGTCTGCTGGCCCTGGGCTTAAGGGCTGGCAGCTCG 2020  
Db 2101 TCCACAGCCAGCAGCAGCTCAGAAAGTCTGCTGGCCCTGGGCTTAAGGGCTGGCAGCTCG 2160  
QY 2021 GATATCTGGCCCGCCAGACCTCTACCCCTTCCCTGACAGCAGCTGTATTTGGCACA 2080  
Db 2161 GATATCTGGCCCGCCAGACCTCTACCCCTTCCCTGACAGCAGCTGTATTTGGCACA 2220  
QY 2081 CAGTCTCTCAGCTTCTACTCTGCTCAGCCATCTACGGAGCGAGTCCAGTTACTCTGCC 2140  
Db 2221 CAGTCTCTCAGCTTCTACTCTGCTCAGCCATCTACGGAGCGAGTCCAGTTACTCTGCC 2280  
QY 2141 TACAGTCCAGCAGCTGCCCACTTGGCGAGACCAAGTCTATTTCTGTGCGCAGCGGCGAG 2200  
Db 2281 TACAGTCCAGCAGCTGCCCACTTGGCGAGACCAAGTCTATTTCTGTGCGCAGCGGCGAG 2340  
QY 2201 AAGCCAGTGCAGAGCTGACTCTCGCGCGAGCTGGCATCAGAGAGCCCTTTGAAAG 2260  
Db 2341 AAGCCAGTGCAGAGCTGACTCTCGCGCGAGCTGGCATCAGAGAGCCCTTTGAAAG 2400  
QY 2261 CAGTTTAAACCGCAGAGCTGCCAAATGGAAATTTGGAGAGCAGTCATGTTCAGAGAACAG 2320  
Db 2401 CAGTTTAAACCGCAGAGCTGCCAAATGGAAATTTGGAGAGCAGTCATGTTCAGAGAACAG 2460  
QY 2321 TCAGGGAGAGCTGGGAAAGTGGGAGCTCAGTCTAGCTTTTCGGGCGAGCATGGAAATC 2380  
Db 2461 TCAGGGAGAGCTGGGAAAGTGGGAGCTCAGTCTAGCTTTTCGGGCGAGCATGGAAATC 2520  
QY 2381 ATTGAGGTCTCTCAGAGAGAAAGACATTTGACTTCTATAGACAAATTTTTTTTCTTCT 2440  
Db 2521 ATTGAGGTCTCTCAGAGAGAAAGACATTTGACTTCTATAGACAAATTTTTTTTCTTCT 2580  
QY 2441 TCACAAAAAATTCCTGTAAATCTGAAATATATATATATATATATATATATATATATAT 2500

Db 2581 TCACAAAAAATTCCTGTGTAATCTGAAATATATATATATATATATATATATATATATAT 2640  
QY 2501 AAAATGGAGCTATGGTGTAAAGCAACAGGTGGATCAACCCAGTTGTACTCTCTTAAACA 2560  
Db 2641 AAAATGGAGCTATGGTGTAAAGCAACAGGTGGATCAACCCAGTTGTACTCTCTTAAACA 2700  
QY 2561 TCTCATTTGAGAGATCAGCTAATCTCTCTCAACAAAATGGAAGGCGAGATGCTAG 2620  
Db 2701 TCTCATTTGAGAGATCAGCTAATCTCTCTCAACAAAATGGAAGGCGAGATGCTAG 2760  
QY 2621 ATCCCCCTTAGAGCGAGGAAAAACATTTTATTCAGTGAATTAACACATCTCTTTGTTCTTA 2680  
Db 2761 ATCCCCCTTAGAGCGAGGAAAAACATTTTATTCAGTGAATTAACACATCTCTTTGTTCTTA 2820  
QY 2681 AAAAAGCAAGTCTTTGGTGTGGAGGACAAAATCCCTACCAATTTTCACAGTTGCT 2740  
Db 2821 AAAAAGCAAGTCTTTGGTGTGGAGGACAAAATCCCTACCAATTTT - CAGGTTGCT 2879  
QY 2741 ACTAAGAGATCTCAAAATATATAGTCTTTGTCGGAGCCCTTCCATAGTACACCTTAGCGCTG 2800  
Db 2880 ACTAAGAGATCTCAAAATATATATATTTAGTCTTTGTCGGAGCCCTTCCATAGTACACCTTAGCGCTG 2939  
QY 2801 AGACTGAGCCAGCTTGGGGTCAAGTAGGTAGAGACCTGTTAGGGACAGAGCCTAGTGTA 2860  
Db 2940 AGACTGAGCCAGCTTGGGGTCAAGTAGGTAGAGACCTGTTAGGGACAGAGCCTAGTGTA 2999  
QY 2861 AATCCAAAGAAATGATCTCTATCCAAAGCTGATTCAAAACCCACGCTCACCTGACAGCC 2920  
Db 3000 AATCCAAAGAAATGATCTCTATCCAAAGCTGATTCAAAACCCACGCTCACCTGACAGCC 3059  
QY 2921 GAGGACACGAGCATCACTCTGTGTCAGGACCAATTAGGGGCTTGCCAAAGGTCTACCTT 2980  
Db 3060 GAGGACACGAGCATCACTCTGTGTCAGGACCAATTAGGGGCTTGCCAAAGGTCTACCTT 3119  
QY 2981 AGACAAACCCAGTACTCTCAGACAGGAAAGTGGGGCTTTGACCCTACCATCTCTGTA 3040  
Db 3120 AGACAAACCCAGTACTCTCAGACAGGAAAGTGGGGCTTTGACCCTACCATCTCTGTA 3179  
QY 3041 GCCCATTTCTAGGCATTTGTGAATAGGTAGTGTAGTACACTTTTCAGACCAATTC 3100  
Db 3180 GCCCATTTCTAGGCATTTGTGAATAGGTAGTGTAGTACACTTTTCAGACCAATTC 3239  
QY 3101 AACGTCTATGCACAAATTCCTGGCTGGCTAGATGAGATAATTTTTTTCTCTCAG 3160  
Db 3240 AACGTCTATGCACAAATTCCTGGCTGGCTAGATGAGATAATTTTTTTCTCTCAG 3299  
QY 3161 CTTTATCAAGAGAGGAAACTCTCTAGATTCAGCTGAAACCAACAGGAACTGGCAACA 3220  
Db 3300 CTTTATCAAGAGAGGAAACTCTCTAGATTCAGCTGAAACCAACAGGAACTGGCAACA 3359  
QY 3221 TCAGATTTAAGCTAAGGTGGAGGCTAACGAGTCTACCTCTTTGTAAATCAAG 3280  
Db 3360 TCAGATTTAAGCTAAGGTGGAGGCTAACGAGTCTACCTCTTTGTAAATCAAG 3419  
QY 3281 AATGTTTAAATGGGATGTCATCTTAAATAAGATGAATTTGGTTTC 3332  
Db 3420 AATGTTTAAATGGGATGTCATCTTAAATAAGATGAATTTGGTTTC 3471

## RESULT 4

ACC60559  
ID ACC60559 standard; cdna; 5450 BP.

XX ACC60559;

XX 19-JUN-2003 (first entry)

DE Polynucleotide relating to the invention SEQ ID NO: 108.

Gene; ss; antiproliferative; hepatotropic; nephrotropic; antiarthritic;  
antiposrotic; cardiac; cyostatic; gene therapy; liver disease;  
proliferative disorder; renal failure; cardiovascular disorder;  
immunological disorder; arthritis; psoriasis; congenital heart defect;



Db 1381 TCTCCAACTTCAATTTCTGGCCCACTCTGGACTATGAGAAGAGATTAGAACCAG 1440  
QY 1301 ACTGGAGCATAGGGCCAAAGAGCAAACTCAAGCTGTGCACTGGAGAAAGCCAAATGAA 1360  
Db 1441 ACTGGAGCATAGGGCCAAAGAGCAAACTCAAGCTGTGCACTGGAGAAAGCCAAATGAA 1500  
QY 1361 CTTGTCCCTGCTGTCTCAGAGGGTGGACAGAAAGAGCAGACGCCCTCAGTCCACCCCTGT 1420  
Db 1501 CTTGTCCCTGCTGTCTCAGAGGGTGGACAGAAAGAGCAGACGCCCTCAGTCCACCCCTGT 1560  
QY 1421 GCGGACTCTGTCTCAGAGGACAGCAAGAAAGAGCCGCTGCAATCCCGCAGCGTCCCC 1480  
Db 1561 GCGGACTCTGTCTCAGAGGACAGCAAGAAAGAGCCGCTGCAATCCCGCAGCGTCCCC 1620  
QY 1481 AGCGTCCAGAGTGCAGCGCTGTCTAGAGGACAGCGCGCTGTACAGGCGCTCAGT 1540  
Db 1621 AGCGTCCAGAGTGCAGCGCTGTCTAGAGGACAGCGCGCTGTACAGGCGCTCAGT 1680  
QY 1541 GGGCTGCACTCTCGGACAGACAGGCTGGAAGACAGCAATAGCTCAAGCGTTCCCTCTCT 1600  
Db 1681 GGGCTGCACTCTCGGACAGACAGGCTGGAAGACAGCAATAGCTCAAGCGTTCCCTCTCT 1740  
QY 1601 CTGGATATCAATCAGTTTCATATTCAGCCAGCATGCGAGCATCTTACATGGCTTCTCC 1660  
Db 1741 CTGGATATCAATCAGTTTCATATTCAGCCAGCATGCGAGCATCTTACATGGCTTCTCC 1800  
QY 1661 TCATCAGAAGATGCTTTTGGAACTACTACAAACCTTCCACTACTCTGGATGGGACCAACAAG 1720  
Db 1801 TCATCAGAAGATGCTTTTGGAACTACTACAAACCTTCCACTACTCTGGATGGGACCAACAAG 1860  
QY 1721 CTATGCCAGTTCTCCCTGTTTCAGAACTATCGAGCAGATCTCCGAAACAGTCTCTGAT 1780  
Db 1861 CTATGCCAGTTCTCCCTGTTTCAGAACTATCGAGCAGATCTCCGAAACAGTCTCTGAT 1920  
QY 1781 AAGGAGGAAGCAGCATCCCAAGAGCTGCAGACCGCAGCGCTTCAGACAGCAGCAGC 1840  
Db 1921 AAGGAGGAAGCAGCATCCCAAGAGCTGCAGACCGCAGCGCTTCAGACAGCAGCAGC 1980  
QY 1841 AAGCATTTGCAATTCGGTTCAGAACTATCGAGCAGATCTCCGAAACAGTCTCTGAT 1900  
Db 1981 AAGCATTTGCAATTCGGTTCAGAACTATCGAGCAGATCTCCGAAACAGTCTCTGAT 2040  
QY 1901 CCACCTGCATCGAAGTGGAGCGTGGAGGACAAATACCAACCCAGCTTCCTTTTCGGCCTT 1960  
Db 2041 CCACCTGCATCGAAGTGGAGCGTGGAGGACAAATACCAACCCAGCTTCCTTTTCGGCCTT 2100  
QY 1961 TCACACGACAGCAGCCTCACGAAGTCTGCTGGCTGGGCTTAAGGCTGGCAGCTCG 2020  
Db 2101 TCACACGACAGCAGCCTCACGAAGTCTGCTGGCTGGGCTTAAGGCTGGCAGCTCG 2160  
QY 2021 GATATCTTGGCCCCCAGACCTCTACCCCTTCCCTGACAGCAGCTGGTATTTGGCACA 2080  
Db 2161 GATATCTTGGCCCCCAGACCTCTACCCCTTCCCTGACAGCAGCTGGTATTTGGCACA 2220  
QY 2081 GAGTCTCAGCACTTCTACTCTGCTCAGCCATCTACGAGGAGCAGTGCAGTACTCTGCC 2140  
Db 2221 GAGTCTCAGCACTTCTACTCTGCTCAGCCATCTACGAGGAGCAGTGCAGTACTCTGCC 2280  
QY 2141 TACAGTCCAGCAGCTGCCCATTTGGAGACCAAGTCTATCTGTCGACGCGGCGAG 2200  
Db 2281 TACAGTCCAGCAGCTGCCCATTTGGAGACCAAGTCTATCTGTCGACGCGGCGAG 2340  
QY 2201 AAGCCAAGTGACAGAGCTGACTCGCGGCGAGTGGCATGAAGAGAGCCCTTTGAAAG 2260  
Db 2341 AAGCCAAGTGACAGAGCTGACTCGCGGCGAGTGGCATGAAGAGAGCCCTTTGAAAG 2400  
QY 2261 CAGTTTAAACGAGAGCTGCCAATTTGGAGAGCATGTCAGAGAAACAG 2320  
Db 2401 CAGTTTAAACGAGAGCTGCCAATTTGGAGAGCATGTCAGAGAAACAG 2460  
QY 2321 TCACGGGAAGAGCTGGGAAAGTGGGAGTCAAGTCTAGCTTTTCGGGACAGTGGAAATC 2380  
Db 2461 TCACGGGAAGAGCTGGGAAAGTGGGAGTCAAGTCTAGCTTTTCGGGACAGTGGAAATC 2520

## RESULT 5

ACC60572

ID ACC60572 standard; cdna; 5450 BP.

XX ACC60572;

AC ACC60572;

XX 19-JUN-2003 (first entry)

DT 19-JUN-2003 (first entry)

QY 2381 ATTGAGGTCTCTCAGAGAAAGACACTTGTGACTTCTATAGACAA-TTTTTTTTCTTG 2439  
Db 2521 ATTGAGGTCTCTCAGAGAAAGACACTTGTGACTTCTATAGACAA-TTTTTTTTCTTG 2580  
QY 2440 TTCACAAAAATTTCCCTGTAAATCTGAATATATATATATATATATATATATATATATTTTG 2499  
Db 2581 TTCACAAAAATTTCCCTGTAAATCTGAATATATATATATATATATATATATATATTTTG 2640  
QY 2500 GAAATGGAGCTATGGTGTAAAGCAACAGGTGATCAACCCAGTTGTTACTCTCTTAAC 2559  
Db 2641 GAAATGGAGCTATGGTGTAAAGCAACAGGTGATCAACCCAGTTGTTACTCTCTTAAC 2700  
QY 2560 ATCTGCATTTTCAGAGATCAGTATATCTTCTCAACAAAATGGAAGGAGAGTGTCTAG 2619  
Db 2701 ATCTGCATTTTCAGAGATCAGTATATCTTCTCTCAACAAAATGGAAGGAGAGTGTCTAG 2760  
QY 2620 AATCCCCCTTAGACGGGAGAAAACCAATTTTATTCAGTGAATTTACATCTCTTGTCTTT 2679  
Db 2761 AATCCCCCTTAGACGGGAGAAAACCAATTTTATTCAGTGAATTTACATCTCTTGTCTTT 2820  
QY 2680 AAAAAAGCAAGTGTCTTTGGTGTGGAGGACAAAATCCCTACATCCATTTTCCAGCTTGTGC 2739  
Db 2821 AAAAAAGCAAGTGTCTTTGGTGTGGAGGACAAAATCCCTACATTTT-CACGTTGTGC 2879  
QY 2740 TACTAAGAGATCTCAAAATATTTAGTCTTTTGTCCGAGCCCTTCCATAGTACACCTTAGCGCT 2799  
Db 2880 TACTAAGAGATCTCAAAATATTTAGTCTTTTGTCCGAGCCCTTCCATAGTACACCTTAGCGCT 2839  
QY 2800 GAGACTGAGCAGCTTGGGGTTCAGGTAGGTAGACCTGTTTAGGACAGAGCTTAGTGT 2859  
Db 2940 GAGACTGAGCAGCTTGGGGTTCAGGTAGGTAGACCTGTTTAGGACAGAGCTTAGTGT 2999  
QY 2860 AAATCCAAAGAAATGATCTTATCCAAAGCTGATTCACAAACCCAGCTCACCTGCAGC 2919  
Db 3000 AAATCCAAAGAAATGATCTTATCCAAAGCTGATTCACAAACCCAGCTCACCTGCAGC 3059  
QY 2920 CGAGGGACACAGAGCATCACTCTGCTGGACGAGCAATTTAGGGGCTTTGCCAAGTCTACCT 2979  
Db 3060 CGAGGGACACAGAGCATCACTCTGCTGGACGAGCAATTTAGGGGCTTTGCCAAGTCTACCT 3119  
QY 2980 TAGAGCAAAACCCAGTACCTCAGACAGGAAGTGGGGCTTTGACCACTACCATATCTGCT 3039  
Db 3120 TAGAGCAAAACCCAGTACCTCAGACAGGAAGTGGGGCTTTGACCACTACCATATCTGCT 3179  
QY 3040 AGCCCATTTTCTAGGCAATTTGAAATAGGTAGGTAGTCAACACTTTTCAGACCAATTC 3099  
Db 3180 AGCCCATTTTCTAGGCAATTTGAAATAGGTAGGTAGTCAACACTTTTCAGACCAATTC 3239  
QY 3100 AAATGCTATGCAAAAATCCGTTGGGCTAGATGGAGATTAATTTTTTTTCTCTCA 3159  
Db 3240 AAATGCTATGCAAAAATCCGTTGGGCTAGATGGAGATTAATTTTTTTTCTCTCA 3299  
QY 3160 GCTTTATGAAGAGAGGAAACTGTCTAGGATTCAGTGAACCAACAGGAACTCTGCAAC 3219  
Db 3300 GCTTTATGAAGAGAGGAAACTGTCTAGGATTCAGTGAACCAACAGGAACTCTGCAAC 3359  
QY 3220 ATCAAGATTTAAGTAAAGTGGAGGCTAAACAGTCTACCTCTCTTTTGTAAATCAA 3279  
Db 3360 ATCAAGATTTAAGTAAAGTGGAGGCTAAACAGTCTACCTCTCTTTTGTAAATCAA 3419  
QY 3280 GAATTTGTTTAAATGGGATTTCAATCTTTTAAATGAAGTGAATTTGTTTC 3332  
Db 3420 GAATTTGTTTAAATGGGATTTCAATCTTTTAAATGAAGTGAATTTGTTTC 3472



QY 1181 AGGATGGACATGTCCTTTAGATGAAGCTTACAGATTTGTGAAGAAAAAGACCTACTATA 1240  
DB 1321 AGGATGGACATGTCCTTTAGATGAAGCTTACAGATTTGTGAAGAAAAAGACCTACTATA 1380  
QY 1241 TCTCCAAACTTCAATTTCTGGGCCAACTCTCGGACATATGAGAAAGATTAAGAACAG 1300  
DB 1381 TCTCCAAACTTCAATTTCTGGGCCAACTCTCGGACATATGAGAAAGATTAAGAACAG 1440  
QY 1301 ACTGGAGCATCAGGGCCAAAGAGCAAACTCAAGCTGTGCACTCTGGAGAGCCAAATGAA 1360  
DB 1441 ACTGGAGCATCAGGGCCAAAGAGCAAACTCAAGCTGTGCACTCTGGAGAGCCAAATGAA 1500  
QY 1361 CCGTCCCTGCTGCTCTCAGAGGGTGGACAGAAAGAGAGAGCCCTCAGTCCACCCCTGT 1420  
DB 1501 CCGTCCCTGCTGCTCTCAGAGGGTGGACAGAAAGAGAGAGCCCTCAGTCCACCCCTGT 1560  
QY 1421 GCGCATCTGTCTCTCAGAGGCGACAGGACAGGACAAAGGCGCGTGCACTCCGCGAGCGTCC 1480  
DB 1561 GCGCATCTGTCTCTCAGAGGCGACAGGACAGGACAAAGGCGCGTGCACTCCGCGAGCGTCC 1620  
QY 1481 AGCGTCCGAGCGTCCGCGTCTGTGTAGAGCAGAGCCGCGTGTACAGGCGCTCAGT 1540  
DB 1621 AGCGTCCGAGCGTCCGCGTCTGTGTAGAGCAGAGCCGCGTGTACAGGCGCTCAGT 1680  
QY 1541 GCGCTGCACCTGTCTCCGACAGAGCGTGGAGACAGCAATAAGCTCAAGCGTTCCTTCTCT 1600  
DB 1681 GCGCTGCACCTGTCTCCGACAGAGCGTGGAGACAGCAATAAGCTCAAGCGTTCCTTCTCT 1740  
QY 1601 CTGGATATCAAACTCAGTTTCATATTCAGCAGAGATGGAGAGTCTTACATGGTCTCTCC 1660  
DB 1741 CTGGATATCAAACTCAGTTTCATATTCAGCAGAGATGGAGAGTCTTACATGGTCTCTCC 1800  
QY 1661 TCATCAGAAAGATGCTTTTGGAACTACTACAAACCTTCCACTACTCTGGATGGAGCAACAAG 1720  
DB 1801 TCATCAGAAAGATGCTTTTGGAACTACTACAAACCTTCCACTACTCTGGATGGAGCAACAAG 1860  
QY 1721 CTATGCCATGTCCTCCCTGTTCAGAACTATTCGAGCAGACTCCCGAAACCACTGCTGAT 1780  
DB 1861 CTATGCCATGTCCTCCCTGTTCAGAACTATTCGAGCAGACTCCCGAAACCACTGCTGAT 1920  
QY 1781 AAGGAGGAAGCAGCATCCCAAGAGCTGCAGACCCGAGGCTTCAGACAGCCAGAGC 1840  
DB 1921 AAGGAGGAAGCAGCATCCCAAGAGCTGCAGACCCGAGGCTTCAGACAGCCAGAGC 1980  
QY 1841 AAGCAATTCGATTTGGTCAGAACAGCAGCAGTGGCAACCGCCAGAGGTCCTTTTATCT 1900  
DB 1981 AAGCAATTCGATTTGGTCAGAACAGCAGCAGTGGCAACCGCCAGAGGTCCTTTTATCT 2040  
QY 1901 CCACCTGCATCGAAGTGGAGCGTGGAGGACAAATACCAACAGCTTCCTTTTCGGCCCTT 1960  
DB 2041 CCACCTGCATCGAAGTGGAGCGTGGAGGACAAATACCAACAGCTTCCTTTTCGGCCCTT 2100  
QY 1961 TCCACAGCCAGCAGCAGCTTCAGAAAGTCTGCTGGCTGGGCTTAAGGGCTGGCACTCG 2020  
DB 2101 TCCACAGCCAGCAGCAGCTTCAGAAAGTCTGCTGGCTGGGCTTAAGGGCTGGCACTCG 2160  
QY 2021 GATATCTTGGCCCCCAGACCTCTACCCCTTCCCTGACAGAGCTGGTATTTTGGCCACA 2080  
DB 2161 GATATCTTGGCCCCCAGACCTCTACCCCTTCCCTGACAGAGCTGGTATTTTGGCCACA 2220  
QY 2081 GAGTCTCTCAGACTTCTACTCTGCTCAGCCATCTACGGAGGAGTGCAGTTACTCTGCC 2140  
DB 2221 GAGTCTCTCAGACTTCTACTCTGCTCAGCCATCTACGGAGGAGTGCAGTTACTCTGCC 2280  
QY 2141 TACAGCTCAGCAGCTGCCACTTTCGGAGAGCAAGTCTATTTCTGTGGCAGGCGGAG 2200  
DB 2281 TACAGCTCAGCAGCTGCCACTTTCGGAGAGCAAGTCTATTTCTGTGGCAGGCGGAG 2340  
QY 2201 AAGCCAGTGCAGAGCTGACTCGCGGCGAGCTGGCATGAGAGAGCCCTTTGGAAG 2260  
DB 2341 AAGCCAGTGCAGAGCTGACTCGCGGCGAGCTGGCATGAGAGAGCCCTTTGGAAG 2400

QY 2261 CAGTTTAAACGACAGAGCTGCCAAATGGAAATTTGGAGAGAGCATCATGTGAGAGACAGG 2320  
DB 2401 CAGTTTAAACGACAGAGCTGCCAAATGGAAATTTGGAGAGAGCATCATGTGAGAGACAGG 2460  
QY 2321 TCACGGGAAGAGCTGGGGAAGTGGGAGTCACTAGCTTTTTCGGGCGACATGGAATC 2380  
DB 2461 TCACGGGAAGAGCTGGGGAAGTGGGAGTCACTAGCTTTTTCGGGCGACATGGAATC 2520  
QY 2381 ATTGAGGCTCTCTCAGAGAAAGACACTTGTGACTTCTATAGACAA-TTTTTTTTCTTG 2439  
DB 2521 ATTGAGGCTCTCTCAGAGAAAGACACTTGTGACTTCTATAGACAAATTTTTTTTCTTG 2580  
QY 2440 TTCAAAAAAATTTCCCTGTAAATCTGAAATATATATATATATATATATATATATTTTG 2499  
DB 2581 TTCAAAAAAATTTCCCTGTAAATCTGAAATATATATATATATATATATATATATTTTG 2640  
QY 2500 GAAATGGAGCTATGTTGTAAAGCAACAGGTGATCAACCCAGTTGTACTCTCTTAAC 2559  
DB 2641 GAAATGGAGCTATGTTGTAAAGCAACAGGTGATCAACCCAGTTGTACTCTCTTAAC 2700  
QY 2560 ATCTGCATTTGAGAGATCAGCTAATCTCTCTCAACAAAAATGGAGGCGACATGCTAG 2619  
DB 2701 ATCTGCATTTGAGAGATCAGCTAATCTCTCTCAACAAAAATGGAGGCGACATGCTAG 2760  
QY 2620 AATCCCCCTAGACGGAGGAAACCAATTTTATTCAGTGAATTACACATCCTCTTGTCTT 2679  
DB 2761 AATCCCCCTAGACGGAGGAAACCAATTTTATTCAGTGAATTACACATCCTCTTGTCTT 2820  
QY 2680 AAAAAGCAAGTGTCTTTGGTGTGGAGGACAAATCCCTACCATTTTCCAGCTTGTGC 2739  
DB 2821 AAAAAGCAAGTGTCTTTGGTGTGGAGGACAAATCCCTACCATTTT-CAGTTGTGC 2879  
QY 2740 TACTAAGAGATCTCAAAATATTAGTCTTTGTCCGAGCCCTTCCATAGTACACCTTAGCGCT 2799  
DB 2880 TACTAAGAGATCTCAAAATATTAGTCTTTGTCCGAGCCCTTCCATAGTACACCTTAGCGCT 2939  
QY 2800 GAGACTGAGCAGCTTGGGGTTCAGGTAGTACCCCTGTTAGGGAAGAGCTTGTGCT 2859  
DB 2940 GAGACTGAGCAGCTTGGGGTTCAGGTAGTACCCCTGTTAGGGAAGAGCTTGTGCT 2999  
QY 2860 AAATCCAAAGAAATGATCCTATCCAAAGCTGATTCACAAACCCAGCTCAGCTGACAGC 2919  
DB 3000 AAATCCAAAGAAATGATCCTATCCAAAGCTGATTCACAAACCCAGCTCAGCTGACAGC 3059  
QY 2920 CGAGGGAACAGAGCATCACTCTCTGTCGACGACCAATTAGGGGCTTTGCCAAGTCTACCT 2979  
DB 3060 CGAGGGAACAGAGCATCACTCTCTGTCGACGACCAATTAGGGGCTTTGCCAAGTCTACCT 3119  
QY 2980 TAGAGCAAAACCCAGTACCTCAGACAGGAAGTGGGGCTTTGACCACTACCATATCTGCT 3039  
DB 3120 TAGAGCAAAACCCAGTACCTCAGACAGGAAGTGGGGCTTTGACCACTACCATATCTGCT 3179  
QY 3040 AGCCCAATTTCTAGGCAATTTGATAGGTAGGTAGTCACTTTTTCAGACCAATTC 3099  
DB 3180 AGCCCAATTTCTAGGCAATTTGATAGGTAGGTAGTCACTTTTTCAGACCAATTC 3239  
QY 3100 AAATGCTATAGCAAAAAATCCCGTGGGCTTAGATGGAGATAATTTTTTTTCTTCTCA 3159  
DB 3240 AAATGCTATAGCAAAAAATCCCGTGGGCTTAGATGGAGATAATTTTTTTTCTTCTCA 3299  
QY 3160 GCTTTATGAAGAGAGGGAACCTGTCTAGATTCAGTCAACCCAGGAACCTGGCAAC 3219  
DB 3300 GCTTTATGAAGAGAGGGAACCTGTCTAGATTCAGTCAACCCAGGAACCTGGCAAC 3359  
QY 3220 ATCAGATTTAAGCTAAGGTGGAGGCTTAAAGAGTCTACCTCCCTCTTTGTAATCAAA 3279  
DB 3360 ATCAGATTTAAGCTAAGGTGGAGGCTTAAAGAGTCTACCTCCCTCTTTGTAATCAAA 3419  
QY 3280 GAATGTTTAAATGGATTTGCAATCCCTTTAAATAAGATGAACCTTGGTTTC 3332  
DB 3420 GAATGTTTAAATGGATTTGCAATCCCTTTAAATAAGATGAACCTTGGTTTC 3472







RESULT 7  
ABK49402  
ID ABK49402 standard; cDNA; 3544 BP.  
XX  
AC ABK49402;  
XX  
DT 02-JUL-2002 (first entry)  
XX  
DE cDNA encoding human dual specificity phosphatase 21117 protein.  
XX  
KW Human; dual specificity phosphatase 21117; erythroid-related disorder;  
KW hematopoietic-related disorder; leukaemia; autoimmune disease; anaemia;  
KW erythrocytosis; liver-related disorder; cancer; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PH Key Location/Qualifiers  
FT CDS 589..2586  
FT /tag= a  
FT /product= "Human dual specificity phosphatase 21117"  
FT /note= "Specifically claimed in claim 2"  
XX  
PN US2002034807-A1.  
XX  
PD 21-MAR-2002.  
XX  
PF 23-MAR-2001; 2001US-00816494.  
XX  
PR 24-MAR-2000; 2000US-0191858P.  
XX  
PA (MEYE/) MEYERS R A.  
XX  
PI Meyers RA;  
XX  
DR WPI; 2002-351088/38.  
DR P-PSDB; AAU79929.  
XX  
PT New nucleic acids, designated 38692 and 21117, encoding dual specificity  
PT phosphatases for treating cell proliferation and differentiation  
PT disorders including hematopoietic and erythroid-related disorders and  
PT cancers.  
XX  
PS Claim 2; Fig 1; 76pp; English.  
XX  
SQ Sequence 3544 BP; 1000 A; 840 C; 822 G; 882 T; 0 U; 0 Other;  
XX  
CC The present invention relates to new nucleic acids designated 38692 and  
CC 21117 encoding dual specificity phosphatase family members. The nucleic  
CC acid, polypeptide encoded by it, and antibody specific for the  
CC polypeptide may be used to diagnose and treat hematopoietic-related  
CC disorders such as leukaemias and autoimmune diseases, erythroid-related  
CC disorders such as anaemias and erythrocytosis, liver-related disorders,  
CC and cancers, particularly of the breast, colon, adipose, prostate and  
CC lung. The present nucleic acid sequence encodes the human dual  
CC specificity phosphatase 21117 protein of the invention, as described  
CC above  
XX  
SQ Sequence 3544 BP; 1000 A; 840 C; 822 G; 882 T; 0 U; 0 Other;  
XX  
CC Query Match 88.5%; Score 2950; DB 6; Length 3544;  
CC Best Local Similarity 95.0%; Pred. No. 0;  
CC Matches 3135; Conservative 0; Mismatches 0; Indels 165; Gaps 2;  
XX  
QY 197 GCTTTTCAGTCCAGTGTAAAGCTGTTGGAGCGGGAGCAAAAGGTAAGATGATGTAATG 256  
Db 224 GCTTTTCAGTCCAGTGTAAAGCTGTTGGAGCGGGAGCAAAAGGTAAGATGATGTAATG 283  
QY 257 CGTGGCTGCTCCAAAGCATCTTTGTTGGAGTGGTATTCAGTCACTCTCTTATGA 316  
Db 284 CGTGGCTGCTCCAAAGCATCTTTGTTGGAGTGGTATTCAGTCACTCTCTTATGA 343  
QY 317 ATCAAAATGTGAGGGGCTGCTTTGTGAGCGGAGTCTTTGCAAGAGCACATCAACGGGAAA 376  
Db 344 ATCAAAATGTGAGGGGCTGCTTTGTGAGCGGAGTCTTTGCAAGAGCACATCAACGGGAAA 403

QY 377 GAGAAAGAGACATTCACTTGGAGGCTCTTCTGCTGAATGGGTTTAACTCTCCTTTTGGC 436  
Db 404 GAGAAAGAGACATTCACTTGGAGGCTCTTCTGCTGAATGGGTTTAACTCTCCTTTTGGC 463  
QY 437 AGTCACCAACAGCCTGACCTCATACATTTTGTAGTCAATGGAGTGGCTGAGCCTTTGAGC 496  
Db 464 AGTCACCAACAGCCTGACCTCATACATTTTGTAGTCAATGGAGTGGCTGAGCCTTTGAGC 523  
QY 497 ACACACCAATTACATCATCGTGGCAATTAAGAGAGGAGTGGGAAAAGAGGACTTATTG 556  
Db 524 ACACACCAATTACATCATCGTGGCAATTAAGAGAGGAGTGGGAAAAGAGGACTTATTG 583  
QY 557 TTGTCATGCCCCATGAGATGATTTGGAATCAAAATGTTTACTGAGAGGTTGGTGGCTCTGC 616  
Db 584 TTGTCATGCCCCATGAGATGATTTGGAATCAAAATGTTTACTGAGAGGTTGGTGGCTCTGC 643  
QY 617 TCGAAAGTGGAAACGGAATAAGTGTCTGCTTAATGTAGCGGCCCATTTGTGGAATACAAATA 676  
Db 644 TCGAAAGTGGAAACGGAATAAGTGTCTGCTTAATGTAGCGGCCCATTTGTGGAATACAAATA 703  
QY 677 CATCCACATTTTGAAGCCATTAATATCAACTGCTCCAAGCTTATGAAGCGAAGGTTGC 736  
Db 704 CATCCACATTTTGAAGCCATTAATATCAACTGCTCCAAGCTTATGAAGCGAAGGTTGC 763  
QY 737 AACAGGACAAAGTGTAAATTACAGAGCTCATCCAGCATTCAGCGAAAACATAAGGTTGACA 796  
Db 764 AACAGGACAAAGTGTAAATTACAGAGCTCATCCAGCATTCAGCGAAAACATAAGGTTGACA 823  
QY 797 TTGATTGCAAGTCCAGAGGTTGTAGTTTACGATCAAGCTCCCAAGATGTTGCCCTCTCTCT 856  
Db 824 TTGATTGCAAGTCCAGAGGTTGTAGTTTACGATCAAGCTCCCAAGATGTTGCCCTCTCTCT 883  
QY 857 CTTGAGACTGTTTCTCACTGTACTTTCTGGGTAAACTGGAGAGAGGTTCAACTCTGTTTC 916  
Db 884 CTTGAGACTGTTTCTCACTGTACTTTCTGGGTAAACTGGAGAGAGGTTCAACTCTGTTTC 943  
QY 917 ACCTGCTTGC----- 926  
Db 944 ACCTGCTTGCAGGTTGGTTTCTGAGTCTCTCGTTGTTTCCCTGGCCTCTGTGAAGGAA 1003  
QY 927 ----- 926  
Db 1004 AATCCACTCTAGTCCCTACCTGCATTTCTCAGCCTTGTCTTACCTGTTGCCAACATTGGGC 1063  
QY 927 -----AGAGC 932  
Db 1064 CAACCCGAATTTTCCCAATCTTTATCTTGCTGCCAGAGATGTCTCTCAACAAGAGC 1123  
QY 933 TGATGCAGCAGAAATGGGATTTGGTTATGTTTAAATGCCAGCAATACCTGTCCAAAGCCTG 992  
Db 1124 TGATGCAGCAGAAATGGGATTTGGTTATGTTTAAATGCCAGCAATACCTGTCCAAAGCCTG 1183  
QY 993 ACTTTATCCCGAGTCTCATTTCTCGGTGTCCTGTGAATGACAGCTTTTGTGAGAAAA 1052  
Db 1184 ACTTTATCCCGAGTCTCATTTCTCGGTGTCCTGTGAATGACAGCTTTTGTGAGAAAA 1243  
QY 1053 TTTTCCGCTGTTGGACAAATCAGTAGATTTCAATTGAGAAAGCAAAAGCCTCCAATGGAT 1112  
Db 1244 TTTTCCGCTGTTGGACAAATCAGTAGATTTCAATTGAGAAAGCAAAAGCCTCCAATGGAT 1303  
QY 1113 GTGTTCTAGTGCATGTTTGTAGTGGATCTCCCGTCCGCCCAACATCGCTATCGCTTACA 1172  
Db 1304 GTGTTCTAGTGCATGTTTGTAGTGGATCTCCCGTCCGCCCAACATCGCTATCGCTTACA 1363  
QY 1173 TCATGAAGGAGTGGACATGCTTTTGTAGTGAAGCTTACAGTTTGTGAAGAAAAAGAC 1232  
Db 1364 TCATGAAGGAGTGGACATGCTTTTGTAGTGAAGCTTACAGTTTGTGAAGAAAAAGAC 1423  
QY 1233 CTACTATATCTCCAAACTTCAATTTTCTGGGCCCAACTCTCTGGACTATGAGAAGAGATTA 1292  
Db 1424 CTACTATATCTCCAAACTTCAATTTTCTGGGCCCAACTCTCTGGACTATGAGAAGAGATTA 1483  
QY 1293 AGAACGAGACTGGAGCATCAGGGCCAAAGCAAACTCAAGCTGCTGCACCTGGAGNAGC 1352



Db	1484	AGAACAGACTGGAGCATCAGGCCCAAGAGCAAACTCAAGCTGCTGCACCTGGAGAGC	1543
Qy	1353	CAAAATGAACCTGTCCCTGTGTCTCAGAGGGTGGACAGAAAGCGAGAGCGCCCTCAGTC	1412
Db	1544	CAAAATGAACCTGTCCCTGTGTCTCAGAGGGTGGACAGAAAGCGAGAGCGCCCTCAGTC	1603
Qy	1413	CACCTGTGCGGACTCTGCTTACCTCAGAGGCAGCAGGACAAAGGCCGTGATCCCGCCA	1472
Db	1604	CACCTGTGCGGACTCTGCTTACCTCAGAGGCAGCAGGACAAAGGCCGTGATCCCGCCA	1663
Qy	1473	GGTGCCCAAGGTGCCCAGCGTCAGCCGTGCTGTTAGAGGACAGCCCGCTGTACAGG	1532
Db	1664	GGTGCCCAAGGTGCCCAGCGTCAGCCGTGCTGTTAGAGGACAGCCCGCTGTGTACAGG	1723
Qy	1533	CGCTCAGTGGGTGCACCTGTCCGAGACAGGCTGGAGACAGCAATAAGCTCAAGCGTT	1592
Db	1724	CGCTCAGTGGGTGCACCTGTCCGAGACAGGCTGGAGACAGCAATAAGCTCAAGCGTT	1783
Qy	1593	CCTTCTCTCTGGATATCAAAATCAGTTTCATATTCAGCCGATGGCAGCATCCTTACATG	1652
Db	1784	CCTTCTCTCTGGATATCAAAATCAGTTTCATATTCAGCCGATGGCAGCATCCTTACATG	1843
Qy	1653	GCTTCTCTCTATCAGAGAGTCTTTGGAATACTACAAACCTTCACCTACTCTGGATGGGA	1712
Db	1844	GCTTCTCTCTATCAGAGAGTCTTTGGAATACTACAAACCTTCACCTACTCTGGATGGGA	1903
Qy	1713	CCAAACAAGCTATGCCAGTTCTCCCTGTTTCCGAGAACTATTCGAGCAGACTCCCGAAACCA	1772
Db	1904	CCAAACAAGCTATGCCAGTTCTCCCTGTTTCCGAGAACTATTCGAGCAGACTCCCGAAACCA	1963
Qy	1773	GTCCTGATTAAGGAGGAAGCCAGCATCCCCAAGAACTGCGACGCGCAGGCGCTTCAGACA	1832
Db	1964	GTCCTGATTAAGGAGGAAGCCAGCATCCCCAAGAACTGCGACGCGCAGGCGCTTCAGACA	2023
Qy	1833	GCCAGAGCAAGCGATTGCAATTCGGTCAGAAACCGAGCAGTGGGACCGCCAGAGTCCC	1892
Db	2024	GCCAGAGCAAGCGATTGCAATTCGGTCAGAAACCGAGCAGTGGGACCGCCAGAGTCCC	2083
Qy	1893	TTTTATCTCCACTGCATCGAAGTGGGAGCGTGGAGGACAATTACCAACACAGCTTCCTTT	1952
Db	2084	TTTTATCTCCACTGCATCGAAGTGGGAGCGTGGAGGACAATTACCAACACAGCTTCCTTT	2143
Qy	1953	TCGGCCTTTCCACGAGCCAGGACCTCAGAAAGTCTGCTGGCCTTGGGCGTTTAAGGGCT	2012
Db	2144	TCGGCCTTTCCACGAGCCAGGACCTCAGAAAGTCTGCTGGCCTTGGGCGTTTAAGGGCT	2203
Qy	2013	GGCACTCGGATATCTTGGCCCCCGAGACCTCTACCCCTTCCCTGACACGAGCTGGATT	2072
Db	2204	GGCACTCGGATATCTTGGCCCCCGAGACCTCTACCCCTTCCCTGACACGAGCTGGATT	2263
Qy	2073	TTGCCACAGAGTCTCTACCTTCTACTCTGCTCAGCCATCTACGAGGACAGTCCAGATT	2132
Db	2264	TTGCCACAGAGTCTCTACCTTCTACTCTGCTCAGCCATCTACGAGGACAGTCCAGATT	2323
Qy	2133	ACTCTGCTACAGCTGCAGCAGCTGCCACTTTCGGAGACCAAGTCTATTCTGTGGCGCA	2192
Db	2324	ACTCTGCTACAGCTGCAGCAGCTGCCACTTTCGGAGACCAAGTCTATTCTGTGGCGCA	2383
Qy	2193	GGCGGAGAGGCAAGTGAACAGAGCTGATCTCGCGCGGAGCTGGGATGAAGAGAGCCCT	2252
Db	2384	GGCGGAGAGGCAAGTGAACAGAGCTGATCTCGCGCGGAGCTGGGATGAAGAGAGCCCT	2443
Qy	2253	TTGAAAGCAGTTTAAACCGAGAGCTGCGCAATTTGGAATTTGGAGAGCAGCATGTGAG	2312
Db	2444	TTGAAAGCAGTTTAAACCGAGAGCTGCGCAATTTGGAATTTGGAGAGCAGCATGTGAG	2503
Qy	2313	AGAACAGGTCACGGGAGAGCTGGGAGAAAGTGGCAGTTCAGTCTAGCTTTTCGGGACGA	2372
Db	2504	AGAACAGGTCACGGGAGAGCTGGGAGAAAGTGGCAGTTCAGTCTAGCTTTTCGGGACGA	2563
Qy	2373	TGGAATCATTTGAGGTCTCTGAGAGAAAGACATTTGTGACTTCTATAGACAATTTTTT	2432

Db	2564	TGGAATCATTTAGGCTCTCTGAGAGAAAGACACTTGTGACTTCTATAGACAAATTTTTT	2622
Qy	2433	TTTCTTTGTTTCAAAAAAATTTCCCTGTAAATCTGAAATATATATATATATGTCATACATATAT	2492
Db	2624	TTTCTTTGTTTCAAAAAAATTTCCCTGTAAATCTGAAATATATATATATGTCATACATATAT	2683
Qy	2493	ATTTTGTGAAATGAGACTATGGTGTAAAAGCAAACAGGTGGATCAACCCAGTTGTTTACTC	2552
Db	2684	ATTTTGTGAAATGAGACTATGGTGTAAAAGCAAACAGGTGGATCAACCCAGTTGTTTACTC	2743
Qy	2553	TCCTTAACATCTGCATTTGAGAGATCAGCTAATACTTCTCTCAACAAAAATGGAAGGCGAG	2612
Db	2744	TCCTTAACATCTGCATTTGAGAGATCAGCTAATACTTCTCTCAACAAAAATGGAAGGCGAG	2803
Qy	2613	ATGCTAGAATCCCCCTAGACGGAGGAAAAACCAATTTTATTCAGTGAATTAACAATCCCTCT	2672
Db	2804	ATGCTAGAATCCCCCTAGACGGAGGAAAAACCAATTTTATTCAGTGAATTAACAATCCCTCT	2863
Qy	2673	TGTTCTTTAAAAAGCAAGTGTCTTTTGGTGTGGAGGCAAAATCCCTACATTTTCCAC	2732
Db	2864	TGTTCTTTAAAAAGCAAGTGTCTTTTGGTGTGGAGGCAAAATCCCTACATTTT - CAC	2922
Qy	2733	GTTGTGCTACTAAGAGATCTCAAAATATTAGTCTTTTGTCCGAGCCCTTCCATAGTACACCT	2792
Db	2923	GTTGTGCTACTAAGAGATCTCAAAATATTAGTCTTTTGTCCGAGCCCTTCCATAGTACACCT	2982
Qy	2793	TAGCGCTGAGACTGAGCCAGCTTTGGGGGTGAGTAGACCCCTGTTAGGGACAGAGCC	2852
Db	2983	TAGCGCTGAGACTGAGCCAGCTTTGGGGGTGAGTAGACCCCTGTTAGGGACAGAGCC	3042
Qy	2853	TAGTGTAAATCCAAAGAAATGATCTCTATCCAAAGCTGATTCACAAACCCAGCTCACC	2912
Db	3043	TAGTGTAAATCCAAAGAAATGATCTCTATCCAAAGCTGATTCACAAACCCAGCTCACC	3102
Qy	2913	TGACAGCCGAGGACACAGCATCACTCTGCTGGACGGACCAATTAGGGGCTTTGCCAAGS	2972
Db	3103	TGACAGCCGAGGACACAGCATCACTCTGCTGGACGGACCAATTAGGGGCTTTGCCAAGS	3162
Qy	2973	TCTACCTTAGAGCAAAACCCAGTACCTCAGACAGGAAAGTCGGGGCTTTGACCACTACCAT	3032
Db	3163	TCTACCTTAGAGCAAAACCCAGTACCTCAGACAGGAAAGTCGGGGCTTTGACCACTACCAT	3222
Qy	3033	ATCTGTAGCCCAATTTTCTAGGCATTTGTAATAGGTAGGTAGTACCTAGTACACTTTTCAGA	3092
Db	3223	ATCTGTAGCCCAATTTTCTAGGCATTTGTAATAGGTAGGTAGTACCTAGTACACTTTTCAGA	3282
Qy	3093	CCAAATCAAACCTGTCATGCAAAAAATTCCTGGGCTTAGATGAGAGATAATTTTTTTTT	3152
Db	3283	CCAAATCAAACCTGTCATGCAAAAAATTCCTGGGCTTAGATGAGAGATAATTTTTTTTT	3342
Qy	3153	CTTCTCAGCTTTATGAAGAGAGGGGAAAATGTCCTAGGATTCAGCTGAAACCCAGGAAACC	3212
Db	3343	CTTCTCAGCTTTATGAAGAGAGGGGAAAATGTCCTAGGATTCAGCTGAAACCCAGGAAACC	3402
Qy	3213	TGGCAATCAGCAATTAAGCTAAGTTGGGAGGTAAACGAGTCTACCTCCCTCTTTGTA	3272
Db	3403	TGGCAATCAGCAATTAAGCTAAGTTGGGAGGTAAACGAGTCTACCTCCCTCTTTGTA	3462
Qy	3273	AATCAAGAAATGTTTAAATGGGATTCATCTCTTAAATAAAGATGAACCTGGTTTC	3332
Db	3463	AATCAAGAAATGTTTAAATGGGATTCATCTCTTAAATAAAGATGAACCTGGTTTC	3522
RESULT 8			
ABV20833			
ID	ABV20833	standard; cDNA; 5145 Bp.	
XX	XX		
AC	ABV20833;		
XX	XX		
DT	13-SEP-2002	(first entry)	
XX	XX		
DE	Human prostate expression marker cDNA 20824.		

POST-T. 8

REV 30933  
RESULT 8

ABV20833  
ID ABV20833 standard; cDNA; 5145 bp.



AC ABV20833;

**XXXXXX**

DT 13-SEP-20

07-SEP-77

11

DE Human pro

XX

KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
KW pharmacogenomic marker; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200160860-A2.  
XX  
XX 23-AUG-2001.  
PD  
XX 20-FEB-2001; 2001WO-US005171.  
XX  
XX 17-FEB-2000; 2000US-0183319P.  
XX  
PR 16-MAR-2000; 2000US-0189862P.  
XX  
PR 25-MAY-2000; 2000US-0207454P.  
XX  
PR 09-JUN-2000; 2000US-0211314P.  
XX  
PR 18-JUL-2000; 2000US-0219007P.  
XX  
PR 13-DEC-2000; 2000US-0255281P.  
XX  
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
XX  
XX Schlegel R, Endege WO, Monahan JE;  
PI  
XX WPI; 2001-662795/76.  
XX  
XX Novel isolated nucleic acid molecule associated with cancerous state of  
PT prostate cells and correlating with presence of prostate cancer, useful  
PT for detecting presence of prostate cancer, stage of prostate cancer.  
XX  
XX Claim 1; Page 3419; 11750pp; English.  
XX  
XX The invention relates to an isolated nucleic acid molecule (I) comprising  
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
CC specification or its complement. (I) is useful for: (a) assessing whether  
CC a patient is afflicted with prostate cancer; (b) monitoring the  
CC progression of prostate cancer in a patient; (c) assessing the efficacy  
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing  
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)  
CC determining whether prostate cancer has metastasized in a patient; (h)  
CC assessing the aggressiveness or indolence of prostate cancer in a patient  
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker  
XX  
XX Sequence 5145 BP; 1408 A; 1135 C; 1253 G; 1346 T; 0 U; 3 Other;  
SQ  
Query Match 88.5%; Score 2950; DB 5; Length 5145;  
Best Local Similarity 95.0%; Pred. No. 0;  
Matches 3135; Conservative 0; Mismatches 0; Indels 165; Gaps 2;  
197 GCTTTCAGTCCAGGTAAAGCTGTTGGAGCGCGGAGCAAGGTAAAGATGATGTAATG 256  
224 GCTTTCAGTCCAGGTAAAGCTGTTGGAGCGCGGAGCAAGGTAAAGATGATGTAATG 293  
257 CGTGGCTGCTCCAAAGCATCTTTTGTGTGGAATGTTATTCAGTCATCTCTTTATGA 316  
284 CGTGGCTGCTCCAAAGCATCTTTTGTGTGGAATGTTATTCAGTCATCTCTTTATGA 343  
317 ATCAATGTGAGGGGCTGTTTGTGAGCGAGTCTTTTGAAGAGCACATCAACGGGAA 376  
344 ATCAATGTGAGGGGCTGTTTGTGAGCGAGTCTTTTGAAGAGCACATCAACGGGAA 403  
377 GAGAAAGAGACATTCATCTTGGAGGGCTCTTGTGAAATGGGTTTAACTCTCTTTTGGC 436  
404 GAGAAAGAGACATTCATCTTGGAGGGCTCTTGTGAAATGGGTTTAACTCTCTTTTGGC 463  
437 AGTCACACAGCCTGACCTCATACATTTTAGTATCAATGAGAGTGGCTTGACCTTTGAGC 496  
464 AGTCACACAGCCTGACCTCATACATTTTAGTATCAATGAGAGTGGCTTGACCTTTGAGC 523  
497 ACACCAACCATTCATCATCTGTCGCAAAATTAAGAGAGGAGGTGGGAAAGAGAGCTTATG 556  
524 ACACCAACCATTCATCATCTGTCGCAAAATTAAGAGAGGAGGTGGGAAAGAGAGCTTATG 593

QY 557 TTGTCATGGCCCATGAGATGATTGGAACTCAAATTTCTTACTGAGAGGTTGGTGGCTCTGC 616  
DB 584 TTGTCATGGCCCATGAGATGATTGGAACTCAAATTTCTTACTGAGAGGTTGGTGGCTCTGC 643  
QY 617 TGGAAAGTGGAAACGGAAGAAAGTCTCTCTAATTATGATAGCGGCCCATTTGTGGAATCAATA 676  
DB 644 TGGAAAGTGGAAACGGAAGAAAGTCTCTCTAATTATGATAGCGGCCCATTTGTGGAATCAATA 703  
QY 677 CATCCACATTTTGGAGCCATTAATCAACTGCTCCCAAGCTTATGAAGCAAGGTTCG 736  
DB 704 CATCCACATTTTGGAGCCATTAATCAACTGCTCCCAAGCTTATGAAGCAAGGTTCG 763  
QY 737 AACAGGACAAAGTGTTAATTACAGAGCTCATCAGCATTCAGGAAACATAAGATTGACA 796  
DB 764 AACAGGACAAAGTGTTAATTACAGAGCTCATCAGCATTCAGGAAACATAAGATTGACA 823  
QY 797 TTGATTCAGTCAGAGGTTGTAGTTTACGATCAAGCTCCCAAGATGTTGCTCTCTCT 856  
DB 824 TTGATTCAGTCAGAGGTTGTAGTTTACGATCAAGCTCCCAAGATGTTGCTCTCTCT 883  
QY 857 CTTTCAGACTGTTTCTCACTGTTCTCTGGTAAACTGGAGAGAGCTTCAACTCTGTTTC 916  
DB 884 CTTTCAGACTGTTTCTCACTGTTCTCTGGTAAACTGGAGAGAGCTTCAACTCTGTTTC 943  
QY 917 ACCTGCTTGC----- 926  
DB 944 ACCTGCTTGCAGGTGGGTTTGTGAGTCTCTCGTTGTTTCCCTGGCCTCTGTGAAGGAA 1003  
QY 927 ----- 926  
DB 1004 AATCCACTAGTCCCTACCTGCATTTCTCAGCCTTGTTCACCATTTGGCAACATTTGGC 1063  
QY 927 -----AGAGC 932  
DB 1064 CAACCCGAATTTTCCCAATCTTTATCTTGGTCGCCAGAGATGCTCCTCAACAGAGAGC 1123  
QY 933 TGATGACAGCAAGTGGATGTTTATGTGTTAAATGCCAGCAATACCTGTCCAAAGCCCTG 992  
DB 1124 TGATGACAGCAAGTGGATGTTTATGTGTTAAATGCCAGCAATACCTGTCCAAAGCCCTG 1183  
QY 993 ACTTTATCCCGAGTCTCACTTCTCGTGTGCTGTGATGACAGCTTGTGGAAGAAA 1052  
DB 1184 ACTTTATCCCGAGTCTCACTTCTCGTGTGCTGTGATGACAGCTTGTGGAAGAAA 1243  
QY 1053 TTTTGGCGTGGTGGCAAAATCAGTAGATTTCAATGAGAAAGCAAAAGCTCCCAATGGAT 1112  
DB 1244 TTTTGGCGTGGTGGCAAAATCAGTAGATTTCAATGAGAAAGCAAAAGCTCCCAATGGAT 1303  
QY 1113 GTGTTCTAGTGCATCTGTTTAGCTGGGATCTCCCGCTCCGCAACCATCGCTATCGCCTACA 1172  
DB 1304 GTGTTCTAGTGCATCTGTTTAGCTGGGATCTCCCGCTCCGCAACCATCGCTATCGCCTACA 1363  
QY 1173 TCATGAGAGAGATGACATGCTCTTTAGATGAAGCTTACAGATTTGTGAAAGAAAAGAGC 1232  
DB 1364 TCATGAGAGAGATGACATGCTCTTTAGATGAAGCTTACAGATTTGTGAAAGAAAAGAGC 1423  
QY 1233 CTACTATATCTCCAAACTCTTCAATTTTCTGGGCCAACTCTCTGGACTATGAGAAAGATTA 1292  
DB 1424 CTACTATATCTCCAAACTCTTCAATTTTCTGGGCCAACTCTCTGGACTATGAGAAAGATTA 1483  
QY 1293 AGAACAGACTGGAGATCAGGAGCAATGCTTGAAGCTTACAGATTTGTGAAAGAAAAGAGC 1352  
DB 1484 AGAACAGACTGGAGATCAGGAGCAATGCTTGAAGCTTACAGATTTGTGAAAGAAAAGAGC 1543  
QY 1353 CAAATGAACCTTCCCTGCTGTCTCAGAGGGTGGACAGAAAAGCGAGACGCCCTCAGTC 1412  
DB 1544 CAAATGAACCTTCCCTGCTGTCTCAGAGGGTGGACAGAAAAGCGAGACGCCCTCAGTC 1603  
QY 1413 CACCTGTGCGCATCTGTCTACCTCAGAGGACAGGACAAAAGCCCGCTGCATCCCGCCA 1472  
DB 1604 CACCTGTGCGCATCTGTCTACCTCAGAGGACAGGACAAAAGCCCGCTGCATCCCGCCA 1663  
QY 1473 GCGTCCCCAGCGTCCCGAGCGTGCAGCCGCTGTTGTAGAGGACAGCCCGCTGCTACAGG 1532



PR 16-MAR-2000; 2000US-0189862P.  
PR 25-MAY-2000; 2000US-0207454P.  
PR 09-JUN-2000; 2000US-0211314P.  
PR 18-JUL-2000; 2000US-0219007P.  
PR 13-DEC-2000; 2000US-0255281P.  
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
XX  
XX Schlegel R, Endege WO, Monahan JB;  
XX WPI; 2001-662795/76.  
XX  
XX Novel isolated nucleic acid molecule associated with cancerous state of  
XX prostate cells and correlating with presence of prostate cancer, useful  
XX for detecting presence of prostate cancer, stage of prostate cancer.  
XX  
XX Claim 1; Page 3481; 11750pp; English.  
XX  
XX The invention relates to an isolated nucleic acid molecule (I) comprising  
XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
XX specification or its complement. (I) is useful for: (a) assessing whether  
XX a patient is afflicted with prostate cancer; (b) monitoring the  
XX progression of prostate cancer in a patient; (c) assessing the efficacy  
XX of a test compound to inhibit prostate cancer in a patient; (d) assessing  
XX the efficacy of a therapy for inhibiting prostate cancer in a patient;  
XX (e) selecting a composition for inhibiting prostate cancer in a patient;  
XX (f) assessing the prostate cell carcinogenic potential of a compound; (g)  
XX determining whether prostate cancer has metastasized in a patient; (h)  
XX assessing the aggressiveness or indolence of prostate cancer in a patient  
XX ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker  
XX  
XX Sequence 5145 BP; 1408 A; 1135 C; 1253 G; 1346 T; 0 U; 3 Other;  
XX  
XX Query Match 88.5%; Score 2950; DB 5; Length 5145;  
XX Best Local Similarity 95.0%; Pred. No. 0;  
XX Matches 3135; Conservative 0; Mismatches 0; Indels 165; Gaps 2;  
XX  
XX 197 GCTTTCAGTCCAGTGAAGCTGTGGAGCGGGAGCAAGGTAAGAAATGATGTAATG 256  
XX 224 GCTTTTCAGTCCAGTGAAGCTGTGGAGCGGGAGCAAGGTAAGAAATGATGTAATG 283  
XX 257 CGTGGCTGCTCCAAAGCATCTTTTGTGGAAATGGTTATCCAGTCACTCTTTTATGA 316  
XX 284 CGTGGCTGCTCCAAAGCATCTTTTGTGGAAATGGTTATCCAGTCACTCTTTTATGA 343  
XX 317 ATCAATATGAGGGGTGCTTTGTGAGCGAGTCTTTGCAAGAGACATCAACGGGAAA 376  
XX 344 ATCAATATGAGGGGTGCTTTGTGAGCGAGTCTTTTGAAGAGACATCAACGGGAAA 403  
XX 377 GAGAAAGAGACATCACTTGGAGGGCTCTTGTGAAATGGTTTAACTCTCTTTTGGCC 436  
XX 404 GAGAAAGAGACATCACTTGGAGGGCTCTTGTGAAATGGTTTAACTCTCTTTTGGCC 463  
XX 437 AGTCACACACAGCTGACCTCATACACTTTTATAGTACAATGGAGTGGCTTGTGAGC 496  
XX 464 AGTCACACACAGCTGACCTCATACACTTTTATAGTACAATGGAGTGGCTTGTGAGC 523  
XX 497 ACACCCACCATTAATCATCGTGGCAATTAAGAGAGGCTGGAAAGAGACTTATTTG 556  
XX 524 ACACCCACCATTAATCATCGTGGCAATTAAGAGAGGCTGGAAAGAGACTTATTTG 583  
XX 557 TTGTTCATGGCCCATGAGATGATTTGAACCTCAATTTTACTGAGAGTTTGTGGCTCTGC 616  
XX 584 TTGTTCATGGCCCATGAGATGATTTGAACCTCAATTTTACTGAGAGTTTGTGGCTCTGC 643  
XX 617 TGGAAAGTGAACGGAAAGAGTCTGCTTAATGATAGCCGGCCATTTGTGGAAATCAATA 676  
XX 644 TGGAAAGTGAACGGAAAGAGTCTGCTTAATGATAGCCGGCCATTTGTGGAAATCAATA 703  
XX 677 CATCCCACTTTTGGAGCCATTAATCAACTGTCTCAAGCTTATGAAGCGAAGGTTGC 736  
XX 704 CATCCCACTTTTGGAGCCATTAATCAACTGTCTCAAGCTTATGAAGCGAAGGTTGC 763  
QY 737 AACAGGACAAAGTGTAAATTACAGAGCTCATCCAGCATTTCCAGCGAAACATAAGTTGACA 796  
DB 764 AACAGGACAAAGTGTAAATTACAGAGCTCATCCAGCATTTCCAGCGAAACATAAGTTGACA 823  
QY 797 TTGATTCAGTCCAGAGGTTGTAGTTTACGATCAAGCTCCCAAGATGTTGCCCTCTCTCT 856  
DB 824 TTGATTCAGTCCAGAGGTTGTAGTTTACGATCAAGCTCCCAAGATGTTGCCCTCTCTCT 883  
QY 857 CTTTCAGACTGTTTCTCACTCTACTCTTCTGGGTAAACTGGAGAGAGCTTCAACTCTCTTC 916  
DB 884 CTTTCAGACTGTTTCTCACTCTACTCTTCTGGGTAAACTGGAGAGAGCTTCAACTCTCTTC 943  
QY 917 ACTCTCTTGC----- 926  
DB 944 ACTCTCTTGCAGGTGGTGTGCTGAGTTCTCTCGTTGTTTCCCTGGCTCTGTGAAGGAA 1003  
QY 927 ----- 926  
DB 1004 AATCCACTCTAGTCCCTACCTGCAATTTCTCAGCCTTGTCTTACCTGTGGCAACATTGGGC 1063  
QY 927 -----AGAGC 932  
DB 1064 CAACCGGAATTTCTCCCAATCTTTTATCTTGCTGCAGCGAGATGCTCTCAACAAGAGC 1123  
QY 933 TGATGACGAGATGGGATTTGTTTATGTTAAATGCCAGCAATACCTGTCCAAAGCCTG 992  
DB 1124 TGATGACGAGATGGGATTTGTTTATGTTAAATGCCAGCAATACCTGTCCAAAGCCTG 1183  
QY 993 ACTTTATCCCCAGTCTCATTTCTTCTGCTGCTGTGAATGACAGCTTTTGTGAGAAA 1052  
DB 1184 ACTTTATCCCCAGTCTCATTTCTGCTGCTGTGATGACAGCTTTTGTGAGAAA 1243  
QY 1053 TTTTGGCGTGTGGACAAATCAGTAGATTTTCATGAGAAAGCAAAAGCTTCCAAATGAT 1112  
DB 1244 TTTTGGCGTGTGGACAAATCAGTAGATTTTCATGAGAAAGCAAAAGCTTCCAAATGAT 1303  
QY 1113 GTGTTCTAGTGACATGTTTAGTGGGATCTCCGCTCGCCACCATCGCTATCGCTTACA 1172  
DB 1304 GTGTTCTAGTGACATGTTTAGTGGGATCTCCGCTCGCCACCATCGCTATCGCTTACA 1363  
QY 1173 TCATGAAGAGATGACATGTTTATGATGAGCTTACAGATTTGTGAAGAAAGAAAGAC 1232  
DB 1364 TCATGAAGAGATGACATGTTTATGATGAGCTTACAGATTTGTGAAGAAAGAAAGAC 1423  
QY 1233 CTACTATATCTCCAACTTCAATTTCTGGGCCAACTCTCGACTATGAGAAAGAAATTA 1292  
DB 1424 CTACTATATCTCCAACTTCAATTTCTGGGCCAACTCTCGACTATGAGAAAGAAATTA 1483  
QY 1293 AGAACCCAGACTGGAGCATCAGGGCCAAAGACAACTCAAGCTCTGCACTGAGAGAGC 1352  
DB 1484 AGAACCCAGACTGGAGCATCAGGGCCAAAGACAACTCAAGCTCTGCACTGAGAGAGC 1543  
QY 1353 CAAATGAACCTGCTGCTGCTCTCAGAGGGTGGACAGAAAGCGAGACGCCCTCAGTC 1412  
DB 1544 CAAATGAACCTGCTGCTGCTCTCAGAGGGTGGACAGAAAGCGAGACGCCCTCAGTC 1603  
QY 1413 CACCTCTGCCACTCTGCTACTCTCAGAGGCGAGGACAAAGGCCCGTGCATCCGCCCA 1472  
DB 1604 CACCTCTGCCACTCTGCTACTCTCAGAGGCGAGGACAAAGGCCCGTGCATCCGCCCA 1663  
QY 1473 GGTGCGCCAGCGTGGCCAGCTGAGCGCTGCTGTTAGAGGACAGCCCGCTGTACAGG 1532  
DB 1664 GGTGCGCCAGCGTGGCCAGCTGAGCGCTGCTGTTAGAGGACAGCCCGCTGTACAGG 1723  
QY 1533 CGTCTAGTGGGTGCACTCTGCTCGGACAGCAGCTGGAGACAGCAATAAGCTCAAGCGTT 1592  
DB 1724 CGTCTAGTGGGTGCACTCTGCTCGGACAGCAGCTGGAGACAGCAATAAGCTCAAGCGTT 1783  
QY 1593 CTTTCTCTCGATATCAAAATCAGTTTCAATTTAGCCAGCATGGCAGCATCTTACATG 1652  
DB 1784 CTTTCTCTCGATATCAAAATCAGTTTCAATTTAGCCAGCATGGCAGCATCTTACATG 1843  
QY 1653 GCTTCTCTCTCATCAGAGATGCTTTGGAAATCTACAAACCTTCCACTACTCTGGATGGGA 1712



PT Novel isolated nucleic acid molecule associated with cancerous state of  
PT prostate cells and correlating with presence of prostate cancer, useful  
XX for detecting presence of prostate cancer, stage of prostate cancer.  
XX  
XX Claim 1; Page 5388-5389; 11750pp; English.

CC The invention relates to an isolated nucleic acid molecule (I) comprising  
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
CC specification or its complement. (I) is useful for: (a) assessing whether  
CC a patient is afflicted with prostate cancer; (b) monitoring the  
CC progression of prostate cancer in a patient; (c) assessing the efficacy  
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing  
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)  
CC determining whether prostate cancer has metastasized in a patient; (h)  
CC assessing the aggressiveness or indolence of prostate cancer in a patient  
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker  
XX  
SQ Sequence 5145 BP; 1408 A; 1135 C; 1253 G; 1346 T; 0 U; 3 Other;

Query Match 88.5%; Score 2950; DB 5; Length 5145;  
Best Local Similarity 95.0%; Pred. No. 0;  
Matches 319; Conservative 0; Mismatches 0; Indels 165; Gaps 2;

QY	197	GCCTTCAGTCCAGTGAAGCTGTTGGAGCGGAGCAAGGTAAGATGATGAATG	256
DB	224	GCCTTCAGTCCAGTGAAGCTGTTGGAGCGGAGCAAGGTAAGATGATGAATG	283
QY	257	CGCTGGCTGCTCAAGAGATCTTTTGTGTGAATGGTTATTCAGTCATCTTTA	316
DB	284	CGCTGGCTGCTCAAGAGATCTTTTGTGTGAATGGTTATTCAGTCATCTTTA	343
QY	317	ATCAAAATGTGAGGGCTGCTTTGTGAGCGAGTCTTTGCAAGAGCACATCAA	376
DB	344	ATCAAAATGTGAGGGCTGCTTTGTGAGCGAGTCTTTGCAAGAGCACATCAA	403
QY	377	GAGAAAGAGACATCTACTTGGAGGGCTCTTGTGAAATGGTTAACTCTCTTT	436
DB	404	GAGAAAGAGACATCTACTTGGAGGGCTCTTGTGAAATGGTTAACTCTCTTT	463
QY	437	AGTCACACAGCTCAGCTCATACACTTTTAGTACAATGGAGTGGCTGAGCCTT	496
DB	464	AGTCACACAGCTCAGCTCATACACTTTTAGTACAATGGAGTGGCTGAGCCTT	523
QY	497	ACACCACTTATCATCATCTGTCGCAATTAAGAGAGAGGTGGGAAAGAGACTT	556
DB	524	ACACCACTTATCATCATCTGTCGCAATTAAGAGAGAGGTGGGAAAGAGACTT	583
QY	557	TTGTCTAGCCCAATGAGATGATGGAACTCAATTTGTTACTGAGAGGTGGTGC	616
DB	584	TTGTCTAGCCCAATGAGATGATGGAACTCAATTTGTTACTGAGAGGTGGTGC	643
QY	617	TGAAAGTGGACGGAAGGCTGCTCAATTTGATAGCGGCAATTTGTGAAATACA	676
DB	644	TGAAAGTGGACGGAAGGCTGCTCAATTTGATAGCGGCAATTTGTGAAATACA	703
QY	677	CATCCCAATTTTGAAGGCAATTAATATCAACTGCTCCAGCTTATGAAGCGAG	736
DB	704	CATCCCAATTTTGAAGGCAATTAATATCAACTGCTCCAGCTTATGAAGCGAG	763
QY	737	AACAGAGCAAGGTGTAATTTACAGAGCTCATCCAGATTCAGCGAAACATAAG	796
DB	764	AACAGAGCAAGGTGTAATTTACAGAGCTCATCCAGATTCAGCGAAACATAAG	823
QY	797	TTGATTGAGTCAAGAGGTGTTAGTTACATCAAGCTCCCAAGATGTTGCCCTCT	856
DB	824	TTGATTGAGTCAAGAGGTGTTAGTTACATCAAGCTCCCAAGATGTTGCCCTCT	883
QY	857	CTTCAGACTGTTTCTCACTGACTCTCGGTAACTGGAGAGAGCTTCACTCTGTC	916
DB	884	CTTCAGACTGTTTCTCACTGACTCTCGGTAACTGGAGAGAGCTTCACTCTGTC	943

QY	917	ACCTGCTTGG-----	926
DB	944	ACCTGCTTGGAGGTGGGTTTCTGCTGAGTTCTCTGTTTCCCTGGGCTCTGTGAAGAA	1003
QY	927	-----	926
DB	1004	AATCCACTCTAGTCCCTACTGCTGCTTCTCAGCCTTGTCTTACCTGTTGCCAACATTTGGC	1063
QY	927	-----	926
DB	1064	CAACCGGAATTTTCCCAATCTTTATCTTGCTGCCAGCAGATGTCTCTCAACAGGAGC	1123
QY	933	TCATGCAGCAAGATGGGATTTGGTTATGTTTAAATGCCAGCAATACCTGTCCTCAAGGCTG	992
DB	1124	TCATGCAGCAAGATGGGATTTGGTTATGTTTAAATGCCAGCAATACCTGTCCTCAAGGCTG	1183
QY	993	ACTTTATCCCGAGTCTCATTTCTCGGTGTGCTGTGAATGACAGCTTTTGTGAGAAA	1052
DB	1184	ACTTTATCCCGAGTCTCATTTCTCGGTGTGCTGTGAATGACAGCTTTTGTGAGAAA	1243
QY	1053	TTTTTCCGTTGGTGGACAAATCAGTAGATTTTCAATTGAGAAAGCAAAAGCCTTCAATGGAT	1112
DB	1244	TTTTTCCGTTGGTGGACAAATCAGTAGATTTTCAATTGAGAAAGCAAAAGCCTTCAATGGAT	1303
QY	1113	GTGTTCTAGTGCATGTTTGTAGTGGATCTCCCGTCCGCCACCATCGCTATCGCTTACA	1172
DB	1304	GTGTTCTAGTGCATGTTTGTAGTGGATCTCCCGTCCGCCACCATCGCTATCGCTTACA	1363
QY	1173	TCATGAAGAGATGACATGCTTTTGTAGTGAAGCTTACAGATTTGTCAAGAAAGAAAGAC	1232
DB	1364	TCATGAAGAGATGACATGCTTTTGTAGTGAAGCTTACAGATTTGTCAAGAAAGAAAGAC	1423
QY	1233	CTACTATATCTCCAACTTCAATTTCTGGGCCAACTCTTGAGCTATGAGAGAGATTA	1292
DB	1424	CTACTATATCTCCAACTTCAATTTCTGGGCCAACTCTTGAGCTATGAGAGAGATTA	1483
QY	1293	AGAAACAGACTGGAGCATCAGGGCCAAAGAGCAAACTCAGCTGCTCAGCTGGAGAGC	1352
DB	1484	AGAAACAGACTGGAGCATCAGGGCCAAAGAGCAAACTCAGCTGCTCAGCTGGAGAGC	1543
QY	1353	CAATGAACCTGTCTCCTGCTCTCAGAGGTGACAGAAAGCGAGACGCCCTCAGTC	1412
DB	1544	CAATGAACCTGTCTCCTGCTCTCAGAGGTGACAGAAAGCGAGACGCCCTCAGTC	1603
QY	1413	CACCTGTGCGCATCTCTCTCAGAGGAGAGAGCAAAAGCCCGTGCATCCGCCCA	1472
DB	1604	CACCTGTGCGCATCTCTCTCAGAGGAGAGAGCAAAAGCCCGTGCATCCGCCCA	1663
QY	1473	GGTGCCAGCTGCGCCAGCTGCGCTGTTAGAGGACAGCCCGTGTGACAGG	1532
DB	1664	GGTGCCAGCTGCGCCAGCTGCGCTGTTAGAGGACAGCCCGTGTGACAGG	1723
QY	1533	CGCTCAGTGGGCTGCACTGTCCGAGACAGGCTGGAGAGAGCAATTAAGCTCAAGCGTT	1592
DB	1724	CGCTCAGTGGGCTGCACTGTCCGAGACAGGCTGGAGAGAGCAATTAAGCTCAAGCGTT	1783
QY	1593	CTTCTCTCTGATATCAATCAGTTTCAATTCAGCCAGCATCGAGCTCCCTTACATG	1652
DB	1784	CTTCTCTCTGATATCAATCAGTTTCAATTCAGCCAGCATCGAGCTCCCTTACATG	1843
QY	1653	GCTTCTCTCTCAGAGATGTTTGGAACTACTAAACCTTCCACTCTCTGAGAGGGA	1712
DB	1844	GCTTCTCTCTCAGAGATGTTTGGAACTACTAAACCTTCCACTCTCTGAGAGGGA	1903
QY	1713	CCAAACAGCTATGCCAGTTCTCCCTGTTTCAAGAACTTATCGGAGCAGACTCCCGAAGCA	1772
DB	1904	CCAAACAGCTATGCCAGTTCTCCCTGTTTCAAGAACTTATCGGAGCAGACTCCCGAAGCA	1963
QY	1773	GTCTCTGATTAAGAGAGAGAGAGAGTCCCAAGAGAGTGCAGACCCGAGGCTTTCAGACA	1832
DB	1964	GTCTCTGATTAAGAGAGAGAGAGAGTCCCAAGAGAGTGCAGACCCGAGGCTTTCAGACA	2023
QY	1833	GCCAGAGACAGGATTTGCTGCTGAGAACAGACAGAGTGGCAGCCGCCAGAGGTCCC	1892

Db 2024 GCAGAGCAGGAGTGCATTCGGTTCAGACAGAGCAGTGGCCACCGCCAGAGGTCCC 2083  
QY 1893 TTTTATCTCCACTGCATCGAAGTGGAGCGTGGAGCAATTAACACACAGAGCTTCCTTT 1952  
Db 2084 TTTTATCTCCACTGCATCGAAGTGGAGCGTGGAGCAATTAACACACAGAGCTTCCTTT 2143  
QY 1953 TCGGCTTTTCCACAGCAGCAGCAGCCTCAGAAAGTCTGCTGGCTGGGCTTTAAGGCT 2012  
Db 2144 TCGGCTTTTCCACAGCAGCAGCAGCCTCAGAAAGTCTGCTGGCTGGGCTTTAAGGCT 2203  
QY 2013 GGCACCTCGATATCTTGGCCCCCAGACCTCTACCCCTTCCCTGACAGCAGCTGGTATT 2072  
Db 2204 GGCACCTCGATATCTTGGCCCCCAGACCTCTACCCCTTCCCTGACAGCAGCTGGTATT 2263  
QY 2073 TTGCCACAGAGTCTTCCACATCTTACTCTGCTCAGCCATCTACGAGGAGTGGCAGTT 2132  
Db 2264 TTGCCACAGAGTCTTCCACATCTTACTCTGCTCAGCCATCTACGAGGAGTGGCAGTT 2323  
QY 2133 ACTCTGCCCTACAGCTGCGAGCAGCTGCCACTTGGGAGACCAAGTCTATTCTGTGGCA 2192  
Db 2324 ACTCTGCCCTACAGCTGCGAGCAGCTGCCACTTGGGAGACCAAGTCTATTCTGTGGCA 2383  
QY 2193 GCGGAGAGCCAAAGTACAGAGCTGACTCGCGGAGAGCTGCGATGAAAGAGAGCCCT 2252  
Db 2384 GCGGAGAGCCAAAGTACAGAGCTGACTCGCGGAGAGCTGCGATGAAAGAGAGCCCT 2443  
QY 2253 TTGAAAGCAGTTTAAACGCGAGAGCTGCCAAATGGAAATTTGGAGAGCATCATGTGAG 2312  
Db 2444 TTGAAAGCAGTTTAAACGCGAGAGCTGCCAAATGGAAATTTGGAGAGCATCATGTGAG 2503  
QY 2313 AGAACAGGTCAACGGAAGAGCTGGGAAAGTGGGAGTCACTGTAGCTTTTCGGGAGCA 2372  
Db 2504 AGAACAGGTCAACGGAAGAGCTGGGAAAGTGGGAGTCACTGTAGCTTTTCGGGAGCA 2563  
QY 2373 TCGAATCATTCAGCTCTCTGAGAGAGACACTTGTGACTCTATACACATATAT 2432  
Db 2564 TCGAATCATTCAGCTCTCTGAGAGAGACACTTGTGACTCTATACACATATAT 2623  
QY 2433 TTCTTGTTTCAAAAAAATCCCTGTAATCTGAAATATATATATATATATATATATAT 2492  
Db 2624 TTCTTGTTTCAAAAAAATCCCTGTAATCTGAAATATATATATATATATATATATAT 2683  
QY 2493 ATTTTGGAAATGAGCTATGGTGTAAAGCAACAGTGGATCAACCCAGTTGTACTC 2552  
Db 2684 ATTTTGGAAATGAGCTATGGTGTAAAGCAACAGTGGATCAACCCAGTTGTACTC 2743  
QY 2553 TCTTAACATCTGCATTTGAGAGATCAGCTAATCTCTCAACAAAAATGGAAGGCGAG 2612  
Db 2744 TCTTAACATCTGCATTTGAGAGATCAGCTAATCTCTCAACAAAAATGGAAGGCGAG 2803  
QY 2613 ATGCTAGAAATCCCTTACAGAGGAGAAACCAATTTTATTCAGTGAATTAACATCCTCT 2672  
Db 2804 ATGCTAGAAATCCCTTACAGAGGAGAAACCAATTTTATTCAGTGAATTAACATCCTCT 2863  
QY 2673 TGTCTTAAAGCAAGTCTTGTGGTGGAGGCAAAATCCCTACCATTTTCCAC 2732  
Db 2864 TGTCTTAAAGCAAGTCTTGTGGTGGAGGCAAAATCCCTACCATTTT-CAC 2922  
QY 2733 GTTGTGCTACTAAGAGATCTCAAAATATTAGTCTTTTGTCCGAGCCCTTCCATAGTACACCT 2792  
Db 2923 GTTGTGCTACTAAGAGATCTCAAAATATTAGTCTTTTGTCCGAGCCCTTCCATAGTACACCT 2982  
QY 2793 TAGCCTGAGACTGAGCAGCTTGGGGTCAAGTATAGTATAGTACCCCTGTAGGACAGAGCC 2852  
Db 2983 TAGCCTGAGACTGAGCAGCTTGGGGTCAAGTATAGTATAGTACCCCTGTAGGACAGAGCC 3042  
QY 2853 TAGTGTAAATCCAAAGAGAAATGATCTTATCAAAAGTGTATTCAAAACCCAGCTCACC 2912  
Db 3043 TAGTGTAAATCCAAAGAGAAATGATCTTATCAAAAGTGTATTCAAAACCCAGCTCACC 3102  
QY 2913 TGACAGCGGAGGACAGCAGATCATCTGCTGGACGGAACATTTAGGGGCTTGGCAAGG 2972

Db 3103 TGACAGCGAGGACACGAGCATCACTCTGCTGACGGAACCATTTAGGGCTTGGCAAGG 3162  
QY 2973 TCTACTTTAGAGCAAAACCCAGTACTCTCAGACAGAAAGTCGGGCTTTTGACCACTACCAT 3032  
Db 3163 TCTACTTTAGAGCAAAACCCAGTACTCTCAGACAGAAAGTCGGGCTTTTGACCACTACCAT 3222  
QY 3033 ATCTGGTAGCCATTTCTTAGGCATTTGTAATAGTAGGTAGCTAGTCACACTTTTTCAGA 3092  
Db 3223 ATCTGGTAGCCATTTCTTAGGCATTTGTAATAGTAGGTAGCTAGTCACACTTTTTCAGA 3282  
QY 3093 CCAATTCAAACTGTCTATGCAAAAATTCCTGGGCTAGATGGAGATAATTTTTTTTT 3152  
Db 3283 CCAATTCAAACTGTCTATGCAAAAATTCCTGGGCTAGATGGAGATAATTTTTTTTT 3342  
QY 3153 CTTCTCAGCTTTTATGAGAGAGGAAAGTCTTAGGATTCAGCTGAACCAACCAAGGAGCC 3212  
Db 3343 CTTCTCAGCTTTTATGAGAGAGGAAAGTCTTAGGATTCAGCTGAACCAACCAAGGAGCC 3402  
QY 3213 TGGCAACATCAGATTTAAGCTAAGGTTGGAGGCTAACGAGTCTACCTCCCTCTTTTGA 3272  
Db 3403 TGGCAACATCAGATTTAAGCTAAGGTTGGAGGCTAACGAGTCTACCTCCCTCTTTTGA 3462  
QY 3273 AATCAAGAAATGTTTAAATGGGATTTCAATCCCTTAAATAAAGATGAACCTGGTTTC 3332  
Db 3463 AATCAAGAAATGTTTAAATGGGATTTCAATCCCTTAAATAAAGATGAACCTGGTTTC 3522  
RESULT 11  
ABV20978  
ID ABV20978 standard; cDNA; 5145 BP.  
XX  
AC ABV20978;  
XX  
DT 13-SEP-2002 (first entry)  
XX  
DE Human prostate expression marker cDNA 20969.  
XX  
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
KW pharmacogenomic marker; gene; ss.  
OS Homo sapiens.  
XX  
XX WO200160860-A2.  
XX  
PD 23-AUG-2001.  
XX  
PF 20-FEB-2001; 2001WO-US005171.  
XX  
PR 17-FEB-2000; 2000US-0183319P.  
PR 16-MAR-2000; 2000US-0189862P.  
PR 25-MAY-2000; 2000US-0207454P.  
PR 09-JUN-2000; 2000US-0211314P.  
PR 18-JUL-2000; 2000US-0219007P.  
PR 13-DEC-2000; 2000US-0255281P.  
XX  
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
XX  
XX Schlegel R, Endege WO, Monahan JE;  
XX  
XX WPI; 2001-662795/76.  
XX  
PT Novel isolated nucleic acid molecule associated with cancerous state of  
PT prostate cells and correlating with presence of prostate cancer, useful  
PT for detecting presence of prostate cancer, stage of prostate cancer.  
XX  
PS Claim 1; Page 3451; 11750pp; English.  
XX  
CC The invention relates to an isolated nucleic acid molecule (I) comprising  
CC a nucleotide sequence given in Tables 1-9 (ABV0010-ABV62213) of the  
CC specification or its complement (II) is useful for: (a) assessing whether  
CC a patient is afflicted with prostate cancer; (b) monitoring the  
CC progression of prostate cancer in a patient; (c) assessing the efficacy  
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing



CC the efficacy of a therapy for inhibiting prostate cancer in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)  
CC determining whether prostate cancer has metastasized in a patient; (h)  
CC assessing the aggressiveness or indolence of prostate cancer in a patient  
CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker  
XX  
SQ Sequence 5145 BP; 1408 A; 1135 C; 1253 G; 1346 T; 0 U; 3 Other;

Query Match		88.5%; Score 2950; DB 5; Length 5145;
Best Local Similarity		95.0%; Pred. No. 0;
Matches 3135; Conservative		0; Mismatches 0; Indels 165; Gaps 2;
QY	197	GCCTTCAGTCCAGTGAAGCTCTCGAGCGGGAGCAAGTAAAGATGATGATG 256
DB	224	GCCTTCAGTCCAGTGAAGCTCTCGAGCGGGAGCAAGTAAAGATGATGATG 283
QY	257	CGCTGGCTGCTCCAAAGCATCTTTTGTGTGGAATGGTTATCCAGTCATCTTTATGA 316
DB	284	CGCTGGCTGCTCCAAAGCATCTTTTGTGTGGAATGGTTATCCAGTCATCTTTATGA 343
QY	317	ATCAAAATGTGAGGGCTGCTTTGTGACGGAGTCCTTTCCAGAGACATCAACGGGAA 376
DB	344	ATCAAAATGTGAGGGCTGCTTTGTGACGGAGTCCTTTCCAGAGACATCAACGGGAA 403
QY	377	GAGAAAGAGACATTCACCTTGGAGGCTCTTGTGAAATGGGTTAACTCTCTTTTGGC 436
DB	404	GAGAAAGAGACATTCACCTTGGAGGCTCTTGTGAAATGGGTTAACTCTCTTTTGGC 463
QY	437	AGTCACACAGCTGACCTACATCTTTTAGTACAAATGGAGTGCTGAGCCTTTGAGC 496
DB	464	AGTCACACAGCTGACCTACATCTTTTAGTACAAATGGAGTGCTGAGCCTTTGAGC 523
QY	497	ACACACCATATACATCATCTGTGCAAAATAAGAGAGGCTGGAAAGAGAGACTTATTG 556
DB	524	ACACACCATATACATCATCTGTGCAAAATAAGAGAGGCTGGAAAGAGAGACTTATTG 583
QY	557	TTGTCATGGCCCATGAGATGATGGAACCAATTTACTGAGAGGTTGGCTCTGTC 616
DB	584	TTGTCATGGCCCATGAGATGATGGAACCAATTTACTGAGAGGTTGGCTCTGTC 643
QY	617	TGGAAGTGAACGGAAAGTCTCTTAATGATAGCCGCCANTTTGGAAATACAATA 676
DB	644	TGGAAGTGAACGGAAAGTCTCTTAATGATAGCCGCCANTTTGGAAATACAATA 703
QY	677	CATCCACATTTTGGAGCCATTAATCACTGCTCCAGCTTATGAGCGAAGGTTGC 736
DB	704	CATCCACATTTTGGAGCCATTAATCACTGCTCCAGCTTATGAGCGAAGGTTGC 763
QY	737	AACAGGACAAAGTTTAATTAAGAGCTCATCCAGCATTCAGGAAACATAAGGTTGACA 796
DB	764	AACAGGACAAAGTTTAATTAAGAGCTCATCCAGCATTCAGGAAACATAAGGTTGACA 823
QY	797	TTGATTCAGTCAGAGGTTGTAGTTTACGATCAAAAGCTCCCAAGATGTTGCTCTCTCT 856
DB	824	TTGATTCAGTCAGAGGTTGTAGTTTACGATCAAAAGCTCCCAAGATGTTGCTCTCTCT 883
QY	857	CTTCAGACTGTTTCTCACTGTAATCTCTGGGTTAACTGGAGAGAGCTTCAACTCTGTC 916
DB	884	CTTCAGACTGTTTCTCACTGTAATCTCTGGGTTAACTGGAGAGAGCTTCAACTCTGTC 943
QY	917	ACCTGCTTTC----- 926
DB	944	ACCTGCTTTCAGGTTGGTTGCTGAGTTCTCTCGTTGTTTCCCTGGCTCTCTGTGAGGAA 1003
QY	927	----- 926
DB	1004	AATCCACTCTAGTCCCTACCTGCATTTCTCAGCCTTGTCTTACCTGTTGCCAACAATTGGGC 1063
QY	927	-----AGGAGC 932
DB	1064	CAACCCGAAATCTTCCCAATCTTTATCTTGGCTGGCCAGCAGATGCTCAACAAGAGGC 1123

QY	933	TGATGACGAGAAATGGGATGGTTATGTGTTAAATGCCAGCAATACCTGTCCAAAGCCTG 992
DB	1124	TGATGACGAGAAATGGGATGGTTATGTGTTAAATGCCAGCAATACCTGTCCAAAGCCTG 1183
QY	993	ACTTTATCCCCAGTCTCATTTCCCTGCGTGTGCTGTGAATGACAGCTTTTGTGAGAAAA 1052
DB	1184	ACTTTATCCCCAGTCTCATTTCCCTGCGTGTGCTGTGAATGACAGCTTTTGTGAGAAAA 1243
QY	1053	TTTTGCGGTGGTTGGACAAATCAGTAGATTTTCATTGAGAAAAAGCAAGCCTCCAAATGGAT 1112
DB	1244	TTTTGCGGTGGTTGGACAAATCAGTAGATTTTCATTGAGAAAAAGCAAGCCTCCAAATGGAT 1303
QY	1113	GTGTTCTAGTGCACATGTTTAGCTGGATCTCCCGTCCGCCACCATCGCTATCGCTTACA 1172
DB	1304	GTGTTCTAGTGCACATGTTTAGCTGGATCTCCCGTCCGCCACCATCGCTATCGCTTACA 1363
QY	1173	TCATGAAGAGATGAGACATGCTCTTTAGATGAAGCTTACAGATTTGTGAAGAAAAAGAC 1232
DB	1364	TCATGAAGAGATGAGACATGCTCTTTAGATGAAGCTTACAGATTTGTGAAGAAAAAGAC 1423
QY	1233	CTACTATATCTCAAACTTCAATTTTCTGGGCCAACTCTCGACTATGAGAGAAAGATTA 1292
DB	1424	CTACTATATCTCAAACTTCAATTTTCTGGGCCAACTCTCGACTATGAGAGAAAGATTA 1483
QY	1293	AGAACAGACTGAGCATCAGGGCCAAAGACAACTCAAGCTGCTGCACCTGGAGAAGC 1352
DB	1484	AGAACAGACTGAGCATCAGGGCCAAAGACAACTCAAGCTGCTGCACCTGGAGAAGC 1543
QY	1353	CAAAATGAACCTGCTCTCTCAGAGGTTGACAGAAAGGAGAGCGCCCTCAGTC 1412
DB	1544	CAAAATGAACCTGCTCTCTCAGAGGTTGACAGAAAGGAGAGCGCCCTCAGTC 1603
QY	1413	CACCTGTGCCAGCTCTGCTACCTCAGAGGAGAGGACAAAGCCCGTGCATCCCGCA 1472
DB	1604	CACCTGTGCCAGCTCTGCTACCTCAGAGGAGAGGACAAAGCCCGTGCATCCCGCA 1663
QY	1473	GGTGGCCAGCTGCGCAGCGTGAGCGCTGCTGTGTTAGAGGACAGCGCGTGTACAGG 1532
DB	1664	GGTGGCCAGCTGCGCAGCGTGAGCGCTGCTGTGTTAGAGGACAGCGCGTGTACAGG 1723
QY	1533	CGCTCAGTGGGTGACCTGCTCCGACAGACAGCTGGAAGACAGCAATAAGCTCAAGCGTT 1592
DB	1724	CGCTCAGTGGGTGACCTGCTCCGACAGACAGCTGGAAGACAGCAATAAGCTCAAGCGTT 1783
QY	1593	CCCTCTCTCTGATATCAAAATCAGTTTCAATTCAGCCAGCATGGCAGCATCTTACATG 1652
DB	1784	CCCTCTCTCTGATATCAAAATCAGTTTCAATTCAGCCAGCATGGCAGCATCTTACATG 1843
QY	1653	GCTTCTCTCTCATCAGAGATGCTTTGGAAATACTACAAACCTTCCACTACTCTGATGGGA 1712
DB	1844	GCTTCTCTCTCATCAGAGATGCTTTGGAAATACTACAAACCTTCCACTACTCTGATGGGA 1903
QY	1713	CCAAAGCTATGCCAGTTTCTCCCTGTTCAGGAACATATCGGAGCAGACTCCCGAAACCA 1772
DB	1904	CCAAAGCTATGCCAGTTTCTCCCTGTTCAGGAACATATCGGAGCAGACTCCCGAAACCA 1963
QY	1773	GTCTCTGATGAAGAGAGAGCCAGCATCCCGACAGAGCTGACAGCCCGCAGGCTTCCAGACA 1832
DB	1964	GTCTCTGATGAAGAGAGAGCCAGCATCCCGACAGAGCTGACAGCCCGCAGGCTTCCAGACA 2023
QY	1833	GCAGAGCAGAGGATTTGCAATTCGGTCAGAACCCAGCAGAGTGGCACCCGCCAGAGGTCCT 1892
DB	2024	GCAGAGCAGAGGATTTGCAATTCGGTCAGAACCCAGCAGAGTGGCACCCGCCAGAGGTCCT 2083
QY	1893	TTTTATCTCACTGCATCGAAGTGGAGGCTGGAGGACAAATTAACAACAGCTTCCCTTT 1952
DB	2084	TTTTATCTCACTGCATCGAAGTGGAGGCTGGAGGACAAATTAACAACAGCTTCCCTTT 2143
QY	1953	TCCGCCCTTTCCACCAAGCAGCAGCACTTCAAGAGTCTGCTGGCTGGGCTTAAAGGCT 2012
DB	2144	TCCGCCCTTTCCACCAAGCAGCAGCACTTCAAGAGTCTGCTGGCTGGGCTTAAAGGCT 2203
QY	2013	GGCACTCGGATATCTTTGGCCCCCAGAGCTCTACCCCTTCCCTGACAGAGCTGGTATT 2072



Db 2204 GGCATCGGATATCTTGGCCCCCAGACCTCTACCCCTTCCTGACGAGCAGTGTAT 2263  
Qy 2073 TTGCCACAGAGTCCTCACATCTTCTGCTCAGCCATCTACGAGGAGTGCAGTT 2132  
Db 2264 TTGCCACAGAGTCCTCACATCTTCTGCTCAGCCATCTACGAGGAGTGCAGTT 2323  
Qy 2133 ACTCTGCTTACAGCTGACGAGCTGCCACTTGGGAGCCAGTCTATCTGTGGCA 2192  
Db 2324 ACTCTGCTTACAGCTGACGAGCTGCCACTTGGGAGCCAGTCTATCTGTGGCA 2383  
Qy 2193 GCGGCGAGAGCCAAAGTACAGAGCTGACTCGCGCGAGCTGCGAGAGAGCCCT 2252  
Db 2384 GCGGCGAGAGCCAAAGTACAGAGCTGACTCGCGCGAGCTGCGAGAGAGCCCT 2443  
Qy 2253 TTGAAAGCAGTTTAAACGCGAGAGCTGCCAAATTTGGAGAGAGCATCATGTGAC 2312  
Db 2444 TTGAAAGCAGTTTAAACGCGAGAGCTGCCAAATTTGGAGAGAGCATCATGTGAC 2503  
Qy 2313 AGAACAGGTCACGGGAAGAGCTGGGGAAGTGGGAGTCAGTCTAGCTTTTCGGGAGCA 2372  
Db 2504 AGAACAGGTCACGGGAAGAGCTGGGGAAGTGGGAGTCAGTCTAGCTTTTCGGGAGCA 2563  
Qy 2373 TGGAAATCATTCAGGTCCTTGAAGAAAGACACTTGTGACTCTATAGACAAATTTTT 2432  
Db 2564 TGGAAATCATTCAGGTCCTTGAAGAAAGACACTTGTGACTCTATAGACAAATTTTT 2623  
Qy 2433 TTTCTTTTTCACAAAAATTCCTGTAAATCTGAATATATATATATATATATATATAT 2492  
Db 2624 TTTCTTTTTCACAAAAATTCCTGTAAATCTGAATATATATATATATATATATATATAT 2583  
Qy 2493 ATTTTGGAAAAATGAGCTATGTTTAAAGCAACAGGTGATCAACCCAGTTGTACTC 2552  
Db 2684 ATTTTGGAAAAATGAGCTATGTTTAAAGCAACAGGTGATCAACCCAGTTGTACTC 2743  
Qy 2553 TCTTAACATCTGCTTTCAGAGATCAGTCTTCTCTCAACAAAAATGGAGGGCAG 2612  
Db 2744 TCTTAACATCTGCTTTCAGAGATCAGTCTTCTCTCAACAAAAATGGAGGGCAG 2803  
Qy 2613 ATGCTAGATCCCCCTAGACGAGAGAAACCAATTTTATTCAGTGAATACACATCTCT 2672  
Db 2804 ATGCTAGATCCCCCTAGACGAGAGAAACCAATTTTATTCAGTGAATACACATCTCT 2863  
Qy 2673 TGTTCTTAAAAAGCAGTGTCTTTGGTTGGAGGACAAATCCCTACCATTTCCAC 2732  
Db 2864 TGTTCTTAAAAAGCAGTGTCTTTGGTTGGAGGACAAATCCCTACCATTTT-CAC 2922  
Qy 2733 GTTGTGCTACTAAGAGATCTCAATATTAGTCTTTGTCCGACCCCTTCCATAGTACACCT 2792  
Db 2923 GTTGTGCTACTAAGAGATCTCAATATTAGTCTTTGTCCGACCCCTTCCATAGTACACCT 2982  
Qy 2793 TAGCGCTGAGACTGACCCAGCTTGGGGTTCAGGTAGGTAGACCTGTGTAGGACAGAGCC 2852  
Db 2983 TAGCGCTGAGACTGACCCAGCTTGGGGTTCAGGTAGGTAGACCTGTGTAGGACAGAGCC 3042  
Qy 2853 TAGTGTGTAATCCAGAGAAATGATCTTATCAAGCTGATTCACAAACCCAGCTCACC 2912  
Db 3043 TAGTGTGTAATCCAGAGAAATGATCTTATCAAGCTGATTCACAAACCCAGCTCACC 3102  
Qy 2913 TGAAGCCGAGGACACGAGCATCTCTGTGTGAGCGACCATTAGGGGCTTCCCAAGG 2972  
Db 3103 TGAAGCCGAGGACACGAGCATCTCTGTGTGAGCGACCATTAGGGGCTTCCCAAGG 3162  
Qy 2973 TCTTACCTTAGACCAACCCAGTACCTCAGACAGGAAGTCGGGCTTTGACCACTACCAT 3032  
Db 3163 TCTTACCTTAGACCAACCCAGTACCTCAGACAGGAAGTCGGGCTTTGACCACTACCAT 3222  
Qy 3033 ATCTGTAGCCCAATTTCTTAGGCATTGTGAATAGGTAGGTAGTACATCACAATTTTCAGA 3092  
Db 3223 ATCTGTAGCCCAATTTCTTAGGCATTGTGAATAGGTAGGTAGTACATCACAATTTTCAGA 3282  
Qy 3093 CCAATTCAACTGTCTATGCCAAAAATTCGGTGGGCTAGATGAGATATTTTTTTTT 3152

Db 3283 CCAATTCAAATGCTATGACAAAAATCCGTGGCCTAGATGGAGATAATTTTTTT 3342  
Qy 3153 CTTCTCAGCTTTATGAAGAGAGAGGAAACTGTCTAGGATTCAGCTGAACCCAGGAACC 3212  
Db 3343 CTTCTCAGCTTTATGAAGAGAGAGGAAACTGTCTAGGATTCAGCTGAACCCAGGAACC 3402  
Qy 3213 TGGCAACATCAGCAATTTAAAGCTAAGGTTGGAGGCTAAACGAGTCTACCTCCCTCTTTGTA 3272  
Db 3403 TGGCAACATCAGCAATTTAAAGCTAAGGTTGGAGGCTAAACGAGTCTACCTCCCTCTTTGTA 3462  
Qy 3273 AATCAAGAAATGTTTAAATGGGATTTGCAATCCTTTAAATAAAGATGAATCTGTTTC 3332  
Db 3463 AATCAAGAAATGTTTAAATGGGATTTGCAATCCTTTAAATAAAGATGAATCTGTTTC 3522

RESULT 12  
ABV21092  
ID ABV21092 standard; cDNA; 5145 BP.  
XX  
AC ABV21092;  
XX  
DT 13-SEP-2002 (first entry)  
XX  
DE Human prostate expression marker cDNA 21083.  
XX  
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
KW pharmacogenomic marker; gene; ss.  
OS Homo sapiens.  
XX  
PN WO200160860-A2.  
XX  
PD 23-AUG-2001.  
XX  
PF 20-FEB-2001; 2001WO-US005171.  
XX  
PR 17-FEB-2000; 2000US-0183319P.  
PR 16-MAR-2000; 2000US-0189862P.  
PR 25-MAY-2000; 2000US-0207454P.  
PR 09-JUN-2000; 2000US-0211314P.  
PR 18-JUL-2000; 2000US-0219007P.  
PR 13-DEC-2000; 2000US-0255281P.  
XX  
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
XX  
PI Schlegel R, Endege WO, Monahan JE;  
XX  
DR WPI; 2001-662795/76.  
XX  
PT Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer.  
XX  
PS Claim 1; Page 3485; 11750pp; English.  
XX  
CC The invention relates to an isolated nucleic acid molecule (I) comprising a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the specification or its complement. (I) is useful for: (a) assessing whether a patient is afflicted with prostate cancer; (b) monitoring the efficacy of progression of prostate cancer in a patient; (c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient; (d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient; (e) selecting a composition for inhibiting prostate cancer in a patient; (f) assessing the prostate cell carcinogenic potential of a compound; (g) determining whether prostate cancer has metastasized in a patient; (h) assessing the aggressiveness or indolence of prostate cancer in a patient; (i) is also useful as a pharmacodynamic or pharmacogenomic marker

Sequence 5145 BP; 1408 A; 1135 C; 1253 G; 1346 T; 0 U; 3 Other;  
Query Match 88.5%; Score 2950; DB 5; Length 5145;  
Best Local Similarity 95.0%; Pred. No. 0;  
Matches 3135; Conservative 0; Mismatches 0; Indels 165; Gaps 2;

QY 197 GCTTTCAGTCCAGTGAAGCTGTTGGAGCGCGGAGCAAAAGGTAAGAATGATGTAATG 256  
DB 224 GCTTTCAGTCCAGTGAAGCTGTTGGAGCGCGGAGCAAAAGGTAAGAATGATGTAATG 283  
QY 257 CGTGGCTGCTCCAAAGCATCTTTTGTGTGGAATGGTTATTCCAGTCACTCTCTTTATGA 316  
DB 284 CGTGGCTGCTCCAAAGCATCTTTTGTGTGGAATGGTTATTCCAGTCACTCTCTTTATGA 343  
QY 317 ATCAAAATGAGGGGCTGCTTTGTGAGCGAGTCTCTTTGCAAGAGCACATCAACGGGAAA 376  
DB 344 ATCAAAATGAGGGGCTGCTTTGTGAGCGAGTCTCTTTGCAAGAGCACATCAACGGGAAA 403  
QY 377 GAGAGAGACATTCACCTTTGGAGGGCTCTTGTGGAATGGTTATTCCAGTCACTCTCTTTGTC 436  
DB 404 GAGAGAGACATTCACCTTTGGAGGGCTCTTGTGGAATGGTTATTCCAGTCACTCTCTTTGTC 463  
QY 437 AGTCACCAACGAGCTGACCTCATACACTTTTGTAGTCAATGAGAGGGCTGAGCCTTTGAGC 496  
DB 464 AGTCACCAACGAGCTGACCTCATACACTTTTGTAGTCAATGAGAGGGCTGAGCCTTTGAGC 523  
QY 497 ACACCAACATTAATCATGTCGCAATTAAGAAGGAGGTGGGAAGAGAGACTTATG 556  
DB 524 ACACCAACATTAATCATGTCGCAATTAAGAAGGAGGTGGGAAGAGAGACTTATG 583  
QY 557 TTGTCTATGCGCCATGAGATGAGTGAACCTCAAAATGTTACTGAGAGGTTGTTGGCTCTGC 616  
DB 584 TTGTCTATGCGCCATGAGATGAGTGAACCTCAAAATGTTACTGAGAGGTTGTTGGCTCTGC 643  
QY 617 TGAAGTGAACGGAAAAGTCTCTTAATGATAGCGGCCATTTGTGGAATACATA 676  
DB 644 TGAAGTGAACGGAAAAGTCTCTTAATGATAGCGGCCATTTGTGGAATACATA 703  
QY 677 CATCCCACTTTTGAAGCATTAAATCAACTGCTCCAAAGCTTATGAAGCGAAGTTGC 736  
DB 704 CATCCCACTTTTGAAGCATTAAATCAACTGCTCCAAAGCTTATGAAGCGAAGTTGC 763  
QY 737 AACAGGCAAAAGTGTAAATPACAGAGCTCATCCAGCAITTCAGCGAAACATAGGTTGACA 796  
DB 764 AACAGGCAAAAGTGTAAATPACAGAGCTCATCCAGCAITTCAGCGAAACATAGGTTGACA 823  
QY 797 TTGATTCAGTCCAGAGGTTGATGTTAGTCAATCAAGCTCCCAAGATGTTGCTCTCT 856  
DB 824 TTGATTCAGTCCAGAGGTTGATGTTAGTCAATCAAGCTCCCAAGATGTTGCTCTCTCT 883  
QY 857 CTTCAGACTGTTTTCTCACTGTACTCTGGGTAAACTGGAGAGAGCTTCAACTCTGTT 916  
DB 884 CTTCAGACTGTTTTCTCACTGTACTCTGGGTAAACTGGAGAGAGCTTCAACTCTGTT 943  
QY 917 ACTGCTTGC----- 926  
DB 944 ACTGCTTGCAGGTGGGTTTGTGAGTCTCTCGTTGTTTCCCTGCGCTCTCTGTAGAGAA 1003  
QY 927 ----- 926  
DB 1004 AATCCACTAGTCCCTACCTGCATTTCTCAGCCTTGCTTACCTGTTGCCAATGTCG 1063  
QY 927 -----AGGAGC 932  
DB 1064 CAACCCGAATTCCTCCCAATCTTTATCTTGGCTGCCAGCGAGATGTCCTCAACAGGAGC 1123  
QY 933 TGATGACAGCAATGGGATGGTTATGTTAAATGCCAGCAATACCTGTCCAAAGCCTG 992  
DB 1124 TGATGACAGCAATGGGATGGTTATGTTAAATGCCAGCAATACCTGTCCAAAGCCTG 1183  
QY 993 ACTTTATCCCGAGTCTCATTTCTCTGCTGCTGTGGAATGACAGCTTTTGTGAGAAA 1052  
DB 1184 ACTTTATCCCGAGTCTCATTTCTCTGCTGCTGTGGAATGACAGCTTTTGTGAGAAA 1243  
QY 1053 TTTTGGCGGTGCGACAAATCAGTATGTTCAATGGAAGCAAAAGCTCCAAATGAT 1112  
DB 1244 TTTTGGCGGTGCGACAAATCAGTATGTTCAATGGAAGCAAAAGCTCCAAATGAT 1303

QY 1113 GTGTTCTAGTGCATGTTTGTAGCTGGATCTCCGCTCCGCCACCATCGCTATCGCTACA 1172  
DB 1304 GTGTTCTAGTGCATGTTTGTAGCTGGATCTCCGCTCCGCCACCATCGCTATCGCTACA 1363  
QY 1173 TCATGAAGAGGATGACATGCTTTTAGATGAAGCTTACAGATTTGTGAAGAAAAGAC 1232  
DB 1364 TCATGAAGAGGATGACATGCTTTTAGATGAAGCTTACAGATTTGTGAAGAAAAGAC 1423  
QY 1233 CTACTATATCTCCAAACTTCAATTTCTGGGCCAACTCTCTGGACTATGAGAGAGATTA 1292  
DB 1424 CTACTATATCTCCAAACTTCAATTTCTGGGCCAACTCTCTGGACTATGAGAGAGATTA 1483  
QY 1293 AGAACACAGACTGGAGCATCAGGGCCAAAGAGCAAACTCAAGCTGCTGACCTGGAGAAGC 1352  
DB 1484 AGAACACAGACTGGAGCATCAGGGCCAAAGAGCAAACTCAAGCTGCTGACCTGGAGAAGC 1543  
QY 1353 CAAATGAACCTGTCCTGCTGCTCAGAGGGTGACAGAAAGCGAGAGCGCCCTCAGTC 1412  
DB 1544 CAAATGAACCTGTCCTGCTGCTCAGAGGGTGACAGAAAGCGAGAGCGCCCTCAGTC 1603  
QY 1413 CACCCTGTGCCGACTCTCTACTCTCAGAGGCGAGGACAAAGGCCGCTGCTATCCGCCA 1472  
DB 1604 CACCCTGTGCCGACTCTCTACTCTCAGAGGCGAGGACAAAGGCCGCTGCTATCCGCCA 1663  
QY 1473 GGTGCCAGCGTGCAGCGTGCAGCGCTGCTGTTAGAGGACAGCGCCGCTGTACAG 1532  
DB 1664 GGTGCCAGCGTGCAGCGTGCAGCGCTGCTGTTAGAGGACAGCGCCGCTGTGTACAG 1723  
QY 1533 CGCTCAGTGGGCTGCACCTGCTCCGAGACAGGCTGGAAGACAGCAATAAGCTCAAGCGTT 1592  
DB 1724 CGCTCAGTGGGCTGCACCTGCTCCGAGACAGGCTGGAAGACAGCAATAAGCTCAAGCGTT 1783  
QY 1593 CTTTCTCTCGATATCAAACTCAGTTTCATATTCAGCCAGCATGGCAGCATCTTACATG 1652  
DB 1784 CTTTCTCTCGATATCAAACTCAGTTTCATATTCAGCCAGCATGGCAGCATCTTACATG 1843  
QY 1653 GTTCTCTCTCATCAGAAGATGCTTTGGAACTACTACAACTTTCCACTACTCTGATGGGA 1712  
DB 1844 GTTCTCTCTCATCAGAAGATGCTTTGGAACTACTACAACTTTCCACTACTCTGATGGGA 1903  
QY 1713 CCAACAAGCTATGCCAGTTTCTCCCTGTTTCAGGAACTATTCGAGAGCAGACTCCCGAAACCA 1772  
DB 1904 CCAACAAGCTATGCCAGTTTCTCCCTGTTTCAGGAACTATTCGAGAGCAGACTCCCGAAACCA 1963  
QY 1773 GTCTGATTAAGAGAGAGCCAGCATCCCAAGAGCTGCAGACCCGACAGGCTTTCAGACA 1832  
DB 1964 GTCTGATTAAGAGAGAGCCAGCATCCCAAGAGCTGCAGACCCGACAGGCTTTCAGACA 2023  
QY 1833 GCCAGAGCAAGGATGTCATTCGCTCAGAACCCAGCAGAGTGGCACCCGCCAGAGGTCCT 1892  
DB 2024 GCCAGAGCAAGGATGTCATTCGCTCAGAACCCAGCAGAGTGGCACCCGCCAGAGGTCCT 2083  
QY 1893 TTTTATCTCCTCATCGATCGAAGTGGAGGGTGGAGGCAATTACCACACAGCTTCCCTTT 1952  
DB 2084 TTTTATCTCCTCATCGATCGAAGTGGAGGGTGGAGGCAATTACCACACAGCTTCCCTTT 2143  
QY 1953 TCGGCTCTTCCACACAGCAGCAGCACCTTCAAGAGTCTGCTGGCTGGGCTTAAAGGCT 2012  
DB 2144 TCGGCTCTTCCACACAGCAGCAGCACCTTCAAGAGTCTGCTGGCTGGGCTTAAAGGCT 2203  
QY 2013 GGCACCTCGGATATCTTGGCCCCCAGACCTCTACCCCTTCCCTGACAGCAGCTGGTATT 2072  
DB 2204 GGCACCTCGGATATCTTGGCCCCCAGACCTCTACCCCTTCCCTGACAGCAGCTGGTATT 2263  
QY 2073 TTGCCACAGAGTCTCCACACTTCTACTCTGCTCAGCCATCTACGAGGAGCAGTGGCAGTT 2132  
DB 2264 TTGCCACAGAGTCTCCACACTTCTACTCTGCTCAGCCATCTACGAGGAGCAGTGGCAGTT 2323  
QY 2133 ACTCTGCTACAGCTGCAGCAGCTGCCCACTTGTGGGAGACCAAGTCTATTCTGTGGCGCA 2192  
DB 2324 ACTCTGCTACAGCTGCAGCAGCTGCCCACTTGTGGGAGACCAAGTCTATTCTGTGGCGCA 2383  
QY 2193 GCGGCGAAGACCAAGCTGACAGAGTGACTCCGCGGGAGCTGGCATGAAGAGAGACCCCT 2252

Db	2384	GGCGCAGAGCCAAAGTGACAGAGCTGACTCGCGCGGAGCTGGCATGAAGAGAGCCCTT	2443
QY	2253	TTGAAAGCAGTTTAAACGCGAAGAGCTGCCAAATGGAATTTGGAGAGAGCATCATGTGACG	2312
Db	2444	TTGAAAGCAGTTTAAACGCGAAGAGCTGCCAAATGGAATTTGGAGAGAGCATCATGTGACG	2503
QY	2313	AGAACAGGTCTACGGGAAAGAGCTGGGAAAGTGGGCAAGTCAGTCTAGCTTTTCGGGCGACGA	2372
Db	2504	AGAACAGGTCTACGGGAAAGAGCTGGGAAAGTGGGCAAGTCAGTCTAGCTTTTCGGGCGACGA	2563
QY	2373	TGGAATCAATTTGAGGTCTCTGAGAGAGAAAGACACATTTGAGCTTCTATAGACAATTTTTT	2432
Db	2564	TGGAATCAATTTGAGGTCTCTGAGAGAGAAAGACACATTTGAGCTTCTATAGACAATTTTTT	2623
QY	2433	TTTCTTGTTCTCAAAAAAAATTCCTCTGTAATCTGAAATATATATATGTACATACATATAT	2492
Db	2624	TTTCTTGTTCTCAAAAAAAATTCCTCTGTAATCTGAAATATATATATGTACATACATATAT	2683
QY	2493	ATTTTGTGAAATCGAGCTATGTTGAAGCAACAGGTGATCAACCCAGGTCTGTACTC	2552
Db	2684	ATTTTGTGAAATCGAGCTATGTTGAAGCAACAGGTGATCAACCCAGGTCTGTACTC	2743
QY	2553	TCTTAACATCTGCATTTTGAGAGATCAGCTAACTTCTCTCAACAAAAATGGAAGGCGAG	2612
Db	2744	TCTTAACATCTGCATTTTGAGAGATCAGCTAACTTCTCTCAACAAAAATGGAAGGCGAG	2803
QY	2613	ATGCTAGAAATCCGCCCTAGACGAGGAAACATTTTATTCAGTGAATTACACATCTCT	2672
Db	2804	ATGCTAGAAATCCGCCCTAGACGAGGAAACATTTTATTCAGTGAATTACACATCTCT	2863
QY	2673	TGTTCTTAAAAAAGCAAGTGTCTTTGGTGTGGAGGACAAAAATCCCTTACCATTTTCCAC	2732
Db	2864	TGTTCTTAAAAAAGCAAGTGTCTTTGGTGTGGAGGACAAAAATCCCTTACCATTTT-CAC	2922
QY	2733	GTTGTGCTACTAAGAGATCTCAAAATATTAGTCTTTGTCGGACCCCTTCCATGATACCT	2792
Db	2923	GTTGTGCTACTAAGAGATCTCAAAATATTAGTCTTTGTCGGACCCCTTCCATGATACCT	2982
QY	2793	TAGCCCTCAGACTCAGACCGACGTTGGGGGTCAGGTAGTACCCCTGTTTGGGACAGAGCC	2852
Db	2983	TAGCCCTCAGACTCAGACCGACGTTGGGGGTCAGGTAGTACCCCTGTTTGGGACAGAGCC	3042
QY	2853	TAGTGGTAAATCCAAGAGAAATGATCCTATCCAAGCTGATTCACAAACCCACGCTCACC	2912
Db	3043	TAGTGGTAAATCCAAGAGAAATGATCCTATCCAAGCTGATTCACAAACCCACGCTCACC	3102
QY	2913	TGACAGCCGAGGACACGAGCATCATCTGCTGGACGACCATTAGGGGGCTTCGCGAAG	2972
Db	3103	TGACAGCCGAGGACACGAGCATCATCTGCTGGACGACCATTAGGGGGCTTCGCGAAG	3162
QY	2973	TCTACCTTTAGAGCAAAACCCAGTACCTCAGACAGAAAGTCGGGGCTTTGACCATACCAT	3032
Db	3163	TCTACCTTTAGAGCAAAACCCAGTACCTCAGACAGAAAGTCGGGGCTTTGACCATACCAT	3222
QY	3033	ATCTGGTAGCCCATTTTCTAGGCATTTGTGAAATAGGTAGGTAGCTAGTCAACTTTTCAGA	3092
Db	3223	ATCTGGTAGCCCATTTTCTAGGCATTTGTGAAATAGGTAGGTAGCTAGTCAACTTTTCAGA	3282
QY	3093	CCAAATTCAACTGTCTATGACACAAAATTCGCCGTGGCCCTAGATGCAGATTAATTTTTTTTT	3152
Db	3283	CCAAATTCAACTGTCTATGACACAAAATTCGCCGTGGCCCTAGATGCAGATTAATTTTTTTTT	3342
QY	3153	CTTCTCAGCTTTATGAAGAGAGGAAACTGTCTAGGATTCAGCTGAACCCACGAGAAC	3212
Db	3343	CTTCTCAGCTTTATGAAGAGAGGAAACTGTCTAGGATTCAGCTGAACCCACGAGAAC	3402
QY	3213	TGGCAACATCAGATTTTAAGCTAAGGTTGGAGGCTTAACGAGCTTACCTCCCTCTTTGTA	3272
Db	3403	TGGCAACATCAGATTTTAAGCTAAGGTTGGAGGCTTAACGAGCTTACCTCCCTCTTTGTA	3462
QY	3273	AATCAAGAAATGTTTAAAAATGGGATTTGTCAATTCCTTTTAAATAAAGATGGAATTTGTTTC	3332

Db	3463	AATCAAGAATAATGTTTAAATGGGATTTGTCATCTTTTAAATGAATGAACATTGGTTTC	3522
RESULT 13			
ABV21312			
ID	ABV21312	standard; cDNA; 5145 BP.	
XX			
AC	ABV21312;		
XX			
DT	13-SEP-2002	(first entry)	
XX			
DE	Human prostate expression marker	cDNA 21303.	
XX			
KW	Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;		
KW	pharmacogenomic marker; gene; ss.		
XX			
OS	Homo sapiens.		
XX			
FN	WO200160860-A2.		
XX			
PD	23-AUG-2001.		
XX			
PF	20-FEB-2001; 2001WO-US005171.		
XX			
PR	17-FEB-2000; 2000US-0183319P.		
PR	16-MAR-2000; 2000US-0189862P.		
PR	25-MAY-2000; 2000US-0207454P.		
PR	09-JUN-2000; 2000US-0211314P.		
PR	18-JUL-2000; 2000US-0219007P.		
PR	13-DEC-2000; 2000US-0255281P.		
XX			
PA	(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.		
XX			
PI	Schlegel R, Endege WO, Monahan JE;		
XX			
XX	WPI; 2001-662795/76.		
DR			
XX			
PT	Novel isolated nucleic acid molecule associated with cancerous state of		
PT	prostate cells and correlating with presence of prostate cancer, useful		
PT	for detecting presence of prostate cancer, stage of prostate cancer.		
XX			
PS	Claim 1; Page 3539; 11750pp; English.		
XX			
CC	The invention relates to an isolated nucleic acid molecule (I) comprising		
CC	a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the		
CC	specification or its complement. (I) is useful for: (a) assessing whether		
CC	a patient is afflicted with prostate cancer; (b) monitoring the		
CC	progression of prostate cancer in a patient; (c) assessing the efficacy		
CC	of a test compound to inhibit prostate cancer in a patient; (d) assessing		
CC	the efficacy of a therapy for inhibiting prostate cancer in a patient;		
CC	(e) selecting a composition for inhibiting prostate cancer in a patient;		
CC	(f) assessing the prostate cell carcinogenic potential of a compound; (g)		
CC	determining whether prostate cancer has metastasized in a patient; (h)		
CC	assessing the aggressiveness or indolence of prostate cancer in a patient		
CC	; (I) is also useful as a pharmacodynamic or pharmacogenomic marker		
XX			
SQ	Sequence 5145 BP; 1408 A; 1135 C; 1253 G; 1346 T; 0 U; 3 Other;		
Query Match	88.5%;	Score 2950; DB 5; Length 5145;	
Best Local Similarity	95.0%;	Pred. No. 0;	
Matches 3135; Conservative	0; Mismatches	0; Indels 165; Gaps	2;
Qy	197	GCCTTCAGTCCAGTGTAAAGCTCTTGAGCGCGGAGCAAGGTAAGAATGATGTAATG	256
Db	224	GCCTTCAGTCCAGTGTAAAGCTCTTGAGCGCGGAGCAAGGTAAGAATGATGTAATG	283
Qy	257	CGCTGGCTGCTCCAAAGCATCTTTTGTGTGGAATGGTATTCCAGTCATCTCTTTATGA	316
Db	284	CGTGGCTGCTCCAAAGCATCTTTTGTGTGGAATGGTATTCCAGTCATCTCTTTATGA	343
Qy	317	ATCAATGTGAGGGGTGCTTTGTGACGGAGTCCCTTTGCAAGAGCATCAACCGGAAA	376
Db	344	ATCAATGTGAGGGGTGCTTTGTGACGGAGTCCCTTTGCAAGAGCATCAACCGGAAA	403

QY 377 GAGAAAGACAAATTCATTGGAGGGCTCTTGCTGAAAATGGGTTTAACCTCTCCTTTTGCC 436  
DB 404 GAGAAAGACAAATTCATTGGAGGGCTCTTGCTGAAAATGGGTTTAACCTCTCCTTTTGCC 463  
QY 437 AGTCACACACAGCTTGACCTCATACACTTTTAGTACAATGAGTGGCTGAGCCTTTGAGC 496  
DB 464 AGTCACACACAGCTTGACCTCATACACTTTTAGTACAATGAGTGGCTGAGCCTTTGAGC 523  
QY 497 ACACCAACCAATCATCATCGTGGGAAAATTAAGAAGGAGGTGGGAAAAGAGGACCTTATTG 556  
DB 524 ACACCAACCAATCATCATCGTGGGAAAATTAAGAAGGAGGTGGGAAAAGAGGACCTTATTG 583  
QY 557 TTGTCAATGCCCATGAGATGATTTGGAATCTCAAAATGTTACTGAGAGGTTGCTGCTCTGC 616  
DB 584 TTGTCAATGCCCATGAGATGATTTGGAATCTCAAAATGTTACTGAGAGGTTGCTGCTCTGC 643  
QY 617 TGGAAAGTGGAAACGGGAAAAGTGTCTTAATTGATAGCGGCCCAATTTGTGGAATACAAATA 676  
DB 644 TGGAAAGTGGAAACGGGAAAAGTGTCTTAATTGATAGCGGCCCAATTTGTGGAATACAAATA 703  
QY 677 CATCCACAAATTTGGAGCCATTAATATCACTGCTCCAGCTTATGAAGCGAAGGTTGC 736  
DB 704 CATCCACAAATTTGGAGCCATTAATATCACTGCTCCAGCTTATGAAGCGAAGGTTGC 763  
QY 737 AACAGACAAAGTGTAAATTACAGAGCTCATCCAGCATTCAGCGAAAACATAAGGTTGACA 796  
DB 764 AACAGACAAAGTGTAAATTACAGAGCTCATCCAGCATTCAGCGAAAACATAAGGTTGACA 823  
QY 797 TTGATTTGAGTCAGAAAGTGTAGTTTACGATCAAGCTCCAGAGTGTGCTCCTCTCT 856  
DB 824 TTGATTTGAGTCAGAAAGTGTAGTTTACGATCAAGCTCCAGAGTGTGCTCCTCTCT 883  
QY 857 CTTTCAGACTGTTTCTTCATCTGACTTCTGGGTAAACTGGAGAAGGCTTCAACTCTGTT 916  
DB 884 CTTTCAGACTGTTTCTTCATCTGACTTCTGGGTAAACTGGAGAAGGCTTCAACTCTGTT 943  
QY 917 ACTGCTTGC----- 926  
DB 944 ACTGCTTGCAGGTGGGTTTGTGAGTTCTCTCGTTGTTTCCCTGGCCTCTGTGAAGGAA 1003  
QY 927 ----- 926  
DB 1004 AATCCACTAGTCCCTACTCGCATTTCTCAGCCTTGCTTACTGTTGTCGAACATTGGGC 1063  
QY 927 -----AGGAGC 932  
DB 1064 CAACCCGAATTTCTCCAAATCTTATCTTGGCTGCCAGCAGATGCTCCTCAACGAAGGC 1123  
QY 933 TGATGACAGCAATGGGATTTGTTATGTTTAAATGCCAGCAATACCTGTCCAAAGCTG 992  
DB 1124 TGATGACAGCAATGGGATTTGTTATGTTTAAATGCCAGCAATACCTGTCCAAAGCTG 1183  
QY 993 ACTTTATCCCCGAGTCTCATTTTCTCGCTGTGCTGTGTAATGACAGCTTTTGTGAGAAA 1052  
DB 1184 ACTTTATCCCCGAGTCTCATTTCTCGCTGTGCTGTGTAATGACAGCTTTTGTGAGAAA 1243  
QY 1053 TTTTGGCTGGTTGGACAAATCAGTAGATTTTCAATGAGNAAGCAAAAGCCTCCATGGAT 1112  
DB 1244 TTTTGGCTGGTTGGACAAATCAGTAGATTTTCAATGAGNAAGCAAAAGCCTCCATGGAT 1303  
QY 1113 GTGTTCTAGTGCACTGTTTAGCTGGGATCTCCCGCTCCGCCACCACTCGCTATCGCCTACA 1172  
DB 1304 GTGTTCTAGTGCACTGTTTAGCTGGGATCTCCCGCTCCGCCACCACTCGCTATCGCCTACA 1363  
QY 1173 TCATGAAGAGGATGGACATGCTCTTTAGATGAAGCTTACAGATTTGTGAAGAAAAGAC 1232  
DB 1364 TCATGAAGAGGATGGACATGCTCTTTAGATGAAGCTTACAGATTTGTGAAGAAAAGAC 1423  
QY 1233 CTACTATCTCCAAATCTCAATTTCTGGGCCAACTCCTCGACTATGAGAAGAGATTA 1292  
DB 1424 CTACTATCTCCAAATCTCAATTTCTGGGCCAACTCCTCGACTATGAGAAGAGATTA 1483

QY 1293 AGAACAGACTGGAGCATCAGGGCCAAAGACAACTCAAGCTGTGCTGACCTGGAGAGC 1352  
DB 1484 AGAACAGACTGGAGCATCAGGGCCAAAGACAACTCAAGCTGTGCTGACCTGGAGAGC 1543  
QY 1353 CAAATGAACCTGTCCTCTGCTGTCTCAGAGGGTGGACAGAAAAGCAGACGCCCTCTCAGTC 1412  
DB 1544 CAAATGAACCTGTCCTCTGCTGTCTCAGAGGGTGGACAGAAAAGCAGACGCCCTCTCAGTC 1603  
QY 1413 CACCCGTGTGGCGACTCTGTACTCTCAGAGGCAGCAGGACAAAGGCCGCTGCATCCGCCCA 1472  
DB 1604 CACCCGTGTGGCGACTCTGTACTCTCAGAGGCAGCAGGACAAAGGCCGCTGCATCCGCCCA 1663  
QY 1473 GGGTCCCGCAGCGTGCAGCGTGCAGCGTGTGTTTAGAGGACAGCCCGCTGGTACAGG 1532  
DB 1664 GGGTCCCGCAGCGTGCAGCGTGCAGCGTGTGTTTAGAGGACAGCCCGCTGGTACAGG 1723  
QY 1533 CGCTCAGTGGGCTGCACCTGTCCGACAGCAGGCTGGAAGACGCAATAAGCTCAAGCGTT 1592  
DB 1724 CGCTCAGTGGGCTGCACCTGTCCGACAGCAGGCTGGAAGACGCAATAAGCTCAAGCGTT 1783  
QY 1593 CCTTCTCTCTGGATATCAAAATCAGTTTCAATTTAGCCAGCATGGCAGCATCTCTTACATG 1652  
DB 1784 CCTTCTCTCTGGATATCAAAATCAGTTTCAATTTAGCCAGCATGGCAGCATCTCTTACATG 1843  
QY 1653 GCTTCTCTCTCAGAGATGCTTTGGAATACCTACAAACCTTCCACTACTCTGGATGGGA 1712  
DB 1844 GCTTCTCTCTCAGAGATGCTTTGGAATACCTACAAACCTTCCACTACTCTGGATGGGA 1903  
QY 1713 CCAACAAAGCTATGCCAGTTCTCCCTGTTCAGGAACTATCGGAGCAGACTCCCGAAACCA 1772  
DB 1904 CCAACAAAGCTATGCCAGTTCTCCCTGTTCAGGAACTATCGGAGCAGACTCCCGAAACCA 1963  
QY 1773 GTCTGTATAAGAGGAAAGCCAGATTCCTCCAAAGAGCTGCGAGACCGCCAGGCTTCAGACA 1832  
DB 1964 GTCTGTATAAGAGGAAAGCCAGATTCCTCCAAAGAGCTGCGAGACCGCCAGGCTTCAGACA 2023  
QY 1833 GCCAGACAAAGGATGTCATTCGCTGAGAACCCAGCAGCAGTGGCACCCGCCAGAGTCCC 1892  
DB 2024 GCCAGACAAAGGATGTCATTCGCTGAGAACCCAGCAGCAGTGGCACCCGCCAGAGTCCC 2083  
QY 1893 TTTTATCTCCACTGTCATCGAAGTGGAGCGTGGAGGACAAATTACCAACAGCTTCCTTT 1952  
DB 2084 TTTTATCTCCACTGTCATCGAAGTGGAGCGTGGAGGACAAATTACCAACAGCTTCCTTT 2143  
QY 1953 TCGGCTTTCCACCAGCAGCAGCAGCCTCAGAGTCTGCTGGCTGGCTGGCTTAAAGGCT 2012  
DB 2144 TCGGCTTTCCACCAGCAGCAGCAGCCTCAGAGTCTGCTGGCTGGCTGGCTTAAAGGCT 2203  
QY 2013 GGCACCTCGGATATCTTGGCCCCCAGACCTCTACCCCTTCCCTGACCAGCAGCTGGTATT 2072  
DB 2204 GGCACCTCGGATATCTTGGCCCCCAGACCTCTACCCCTTCCCTGACCAGCAGCTGGTATT 2263  
QY 2073 TTGCCACAGAGTCTCAGCTTCTACTTGCTCAGCCATCTACGGAGGAGTGCAGTT 2132  
DB 2264 TTGCCACAGAGTCTCAGCTTCTACTTGCTCAGCCATCTACGGAGGAGTGCAGTT 2323  
QY 2133 ACTCTGCTTACAGTGCAGCAGCTGCCACTTGGGAGACCAAGTCTATTCTGTGGCA 2192  
DB 2324 ACTCTGCTTACAGTGCAGCAGCTGCCACTTGGGAGACCAAGTCTATTCTGTGGCA 2383  
QY 2193 GCGCGCAGAGCAAGTGCAGAGCTGACTCGCGCGGAGCTGGCATGAGAGAGCCCT 2252  
DB 2384 GCGCGCAGAGCAAGTGCAGAGCTGACTCGCGCGGAGCTGGCATGAGAGAGCCCT 2443  
QY 2253 TTGAAAGCAGTTTAAACGCGAGAAGCTGCCAAATGGAATTTGGAGAGAGCATCATGTGAG 2312  
DB 2444 TTGAAAGCAGTTTAAACGCGAGAAGCTGCCAAATGGAATTTGGAGAGAGCATCATGTGAG 2503  
QY 2313 AGHACAGGTACCGGAAAGAGCTGGGAAAGTGGGAGTCAGTCTAGCTTTTCGGGGACGA 2372  
DB 2504 AGHACAGGTACCGGAAAGAGCTGGGAAAGTGGGAGTCAGTCTAGCTTTTCGGGGACGA 2563  
QY 2373 TGGAAATCATTTAGAGTCTCTCTGAGAGGAAAGACACTTGTGACTTCTATAGACAAATTTTT 2432



QY 557 TTGTCATGCCCATGAGATGATTGGAATCTCAAATGTTTACTGAGAGGTTGGTGCCTCTGC 616  
DB 584 TTGTCATGCCCATGAGATGATTGGAATCTCAAATGTTTACTGAGAGGTTGGTGCCTCTGC 643  
QY 617 TGGAAAGTGGAAAGGAAAGTGGCTTAATTTGATAGCGGCCCATTTGTTGGAATACAAATA 676  
DB 644 TGGAAAGTGGAAAGGAAAGTGGCTTAATTTGATAGCGGCCCATTTGTTGGAATACAAATA 703  
QY 677 CATCCACACATTTTGAAGGCATTAATATCAACTGCTCCAAGCTTATGAAGCGAAAGTTGC 736  
DB 704 CATCCACACATTTTGAAGGCATTAATATCAACTGCTCCAAGCTTATGAAGCGAAAGTTGC 763  
QY 737 AACAGACAAAGTGTAAATACAGAGCTCATCCAGCATTCAGCGAAACATGAAGTTGACA 796  
DB 764 AACAGACAAAGTGTAAATACAGAGCTCATCCAGCATTCAGCGAAACATGAAGTTGACA 823  
QY 797 TTGATTTGAGTCAAGAGGTTGATTTTACGATCAAAAGCTCCCAAGATGTTGCCTCTCTCT 856  
DB 824 TTGATTTGAGTCAAGAGGTTGATTTTACGATCAAAAGCTCCCAAGATGTTGCCTCTCTCT 883  
QY 857 CTTTACAGACTGTTTCTCACTGATCTTCTGGGTAAACTGGAGAGAGCTTCAACTCTGTTT 916  
DB 884 CTTTACAGACTGTTTCTCACTGATCTTCTGGGTAAACTGGAGAGAGCTTCAACTCTGTTT 943  
QY 917 ACCTGCTTGC----- 926  
DB 944 ACCTGCTTGCAGTGGGTTTGTGAGTCTCTCGTTGTTTCCCTGGCCTCTGTGAAGAA 1003  
QY 927 ----- 926  
DB 1004 AATCCACTAGTCCCTACCTGCATTTCTCAGCCTTGTCTTACCTGTGCAACATTTGGGC 1063  
QY 927 -----AGGAGC 932  
DB 1064 CAACCCGAAATCTTCCCAATCTTTATCTTGTGGTCCGACGAGATGCTCTCAACAGGAGC 1123  
QY 933 TGATGCAGCAGAATGGGATTTGTTATGTTTAAATGCCAGCAATACCTGTCCAAAGCCTG 992  
DB 1124 TGATGCAGCAGAATGGGATTTGTTATGTTTAAATGCCAGCAATACCTGTCCAAAGCCTG 1183  
QY 993 ACTTTATCCCGAGTCTCATTTCTCGGTGTGCTGTGAATGACAGCTTTTGTGAGAAA 1052  
DB 1184 ACTTTATCCCGAGTCTCATTTCTCGGTGTGCTGTGAATGACAGCTTTTGTGAGAAA 1243  
QY 1053 TTTTGGCGTGGTTGGACAAATCAGTAGATTTCAATGAGAAAGCAAAAGCCTCCAAATGGAT 1112  
DB 1244 TTTTGGCGTGGTTGGACAAATCAGTAGATTTCAATGAGAAAGCAAAAGCCTCCAAATGGAT 1303  
QY 1113 GTGTTCTAGTGCACTGTTTATGCTGGGATCTCCGCTCCGACCATCGCTATCGCTTACA 1172  
DB 1304 GTGTTCTAGTGCACTGTTTATGCTGGGATCTCCGCTCCGACCATCGCTATCGCTTACA 1363  
QY 1173 TCATGAGAGGATGACATGCTTTTATGATGAAGCTTACAGATTTGTGAAGAAAAAGAC 1232  
DB 1364 TCATGAGAGGATGACATGCTTTTATGATGAAGCTTACAGATTTGTGAAGAAAAAGAC 1423  
QY 1233 CTACTATATCTCCAAACTTCAATTTTCTGGGCCAACTCTTGGACTATGAGAGAGATTA 1292  
DB 1424 CTACTATATCTCCAAACTTCAATTTTCTGGGCCAACTCTTGGACTATGAGAGAGATTA 1483  
QY 1293 AGAACAGACTGGAGCATCAGGCGCAAGAGCAAACTCAAGCTGTGCACTTGGAGAGC 1352  
DB 1484 AGAACAGACTGGAGCATCAGGCGCAAGAGCAAACTCAAGCTGTGCACTTGGAGAGC 1543  
QY 1353 CAAATGAACCTGTCCCTGCTGTCTCAGAGGTTGACAGAGAAAGCGAGACGCCCTCAGTTC 1412  
DB 1544 CAAATGAACCTGTCCCTGCTGTCTCAGAGGTTGACAGAGAAAGCGAGACGCCCTCAGTTC 1603  
QY 1413 CACCTGTGCGACTCTGCTACTCTCAGAGGAGCAGGACAAAGCGCGTGCATCCGCCCA 1472  
DB 1604 CACCTGTGCGACTCTGCTACTCTCAGAGGAGCAGGACAAAGCGCGTGCATCCGCCCA 1663

QY 1473 GCGTGCAGAGCTGCCCGCAGCGTGCAGCGCTGCTGTTAGAGGACAGCCGCTGTTACAGG 1532  
DB 1664 GCGTGCAGAGCTGCCCGCAGCGTGCAGCGCTGCTGTTAGAGGACAGCCGCTGTTACAGG 1723  
QY 1533 GCGTCAGTGGGCTGCACCTGTCCGACAGACAGGCTGGAAGACAGCAATAAGCTCAAGCGTT 1592  
DB 1724 GCGTCAGTGGGCTGCACCTGTCCGACAGACAGGCTGGAAGACAGCAATAAGCTCAAGCGTT 1783  
QY 1593 CTTTCTCTCTGGATATCAAAATCAGTTTCAATTTAGAGGACAGCTTCCGAAACCA 1652  
DB 1784 CTTTCTCTCTGGATATCAAAATCAGTTTCAATTTAGAGGACAGCTTCCGAAACCA 1843  
QY 1653 GCTTCTCTCTCATCAGAGATGCTTTTGAATACTACAAACCTTCCACTACTCTGGATGGGA 1712  
DB 1844 GCTTCTCTCTCATCAGAGATGCTTTTGAATACTACAAACCTTCCACTACTCTGGATGGGA 1903  
QY 1713 CCAAACAGCTATGCCAGTTCTCCCTGTTTCAAGAACTTATCGGAGCAGACTCCCGAAACCA 1772  
DB 1904 CCAAACAGCTATGCCAGTTCTCCCTGTTTCAAGAACTTATCGGAGCAGACTCCCGAAACCA 1963  
QY 1773 GTCTGTAAAGGAGGAGGAGCCAGCATCCCAAGAAAGCTGCAGACCCGAGGCTTCAGACA 1832  
DB 1964 GTCTGTAAAGGAGGAGGAGCCAGCATCCCAAGAAAGCTGCAGACCCGAGGCTTCAGACA 2023  
QY 1833 GCGAGACAAAGGATTTGATTTGCTGTCAGAACCCAGCAGCTGGCAGCCGAGAGGTCCTC 1892  
DB 2024 GCGAGACAAAGGATTTGATTTGCTGTCAGAACCCAGCAGCTGGCAGCCGAGAGGTCCTC 2083  
QY 1893 TTTTATCTCACTGATCGAAAGTGGAGCGGTGGAGGCAATACCAACACAGCTTCCTTT 1952  
DB 2084 TTTTATCTCACTGATCGAAAGTGGAGCGGTGGAGGCAATACCAACACAGCTTCCTTT 2143  
QY 1953 TCGGCTTTTCCACAGGACAGCAGCACTCTAGAAAGTCTGCTGGCTGGGCTTAAAGGCT 2012  
DB 2144 TCGGCTTTTCCACAGGACAGCAGCACTCTAGAAAGTCTGCTGGCTGGGCTTAAAGGCT 2203  
QY 2013 GGCACCTCGATATCTTGGCCCCCAGAGCTCTACCCCTTCCCTGACACAGCAGCTGTAAT 2072  
DB 2204 GGCACCTCGATATCTTGGCCCCCAGAGCTCTACCCCTTCCCTGACACAGCAGCTGTAAT 2263  
QY 2073 TTGCCACAGAGTCTCACACTTCTACTCTGCTCAGCCATCTACGGAGGAGTGCAGTT 2132  
DB 2264 TTGCCACAGAGTCTCACACTTCTACTCTGCTCAGCCATCTACGGAGGAGTGCAGTT 2323  
QY 2133 ACTTGCCTACAGCTGCAGCAGCTGCCACTTGGGAGAGCAAGTCTATTCTGTGCGCA 2192  
DB 2324 ACTTGCCTACAGCTGCAGCAGCTGCCACTTGGGAGAGCAAGTCTATTCTGTGCGCA 2383  
QY 2193 GCGCGCAGAGCCAAAGTGCAGAGCTGACTCGCGCGGAGCTGGCATGAAGAGAGCCCT 2252  
DB 2384 GCGCGCAGAGCCAAAGTGCAGAGCTGACTCGCGCGGAGCTGGCATGAAGAGAGCCCT 2443  
QY 2253 TTGAAAGCGATTTAAACCGCAGAGCTGCCAAATGGAATTTGGAGAGAGCATCATGTGAG 2312  
DB 2444 TTGAAAGCGATTTAAACCGCAGAGCTGCCAAATGGAATTTGGAGAGAGCATCATGTGAG 2503  
QY 2313 AGAACAGCTCAAGGAGAGCTGGGAAAGTGGGAGTGCAGTCTAGCTTTTGGGCGAGCA 2372  
DB 2504 AGAACAGCTCAAGGAGAGCTGGGAAAGTGGGAGTGCAGTCTAGCTTTTGGGCGAGCA 2563  
QY 2373 TGGAAATCATTTGAGGCTCTCTGAGAGAGAGACACTTGTGACTTCTATAGACAAATTTT 2432  
DB 2564 TGGAAATCATTTGAGGCTCTCTGAGAGAGAGACACTTGTGACTTCTATAGACAAATTTT 2623  
QY 2433 TTTTCTTCTTCAAAAAAATTTCCCTGTAATCTGAAATATATATATATATATATATATAT 2492  
DB 2624 TTTTCTTCTTCAAAAAAATTTCCCTGTAATCTGAAATATATATATATATATATATATAT 2683  
QY 2493 ATTTTTCGAAAAATGAGAGCTATGTTGTAAGGCAACAGGTGGATCAACCCAGTTGTTACTC 2552  
DB 2684 ATTTTTCGAAAAATGAGAGCTATGTTGTAAGGCAACAGGTGGATCAACCCAGTTGTTACTC 2743  
QY 2553 TCTTAAACATCTGCATTTGAGAGATCAGTAACTCTCTCAACAAAAATGGAAGGCGAG 2612







Db	2923	GTGTGCTACTAAGAGATCTCAATATATAGTCTTTGTCGGACCCCTTCATATGATACACT	2982
Qy	2793	TAGCGCTGAGACTGAGCCAGCTTTGGGGGTGAGTAGTACCCCTGTTTGGGACAGAGCC	2852
Db	2983	TAGCGCTGAGACTGAGCCAGCTTTGGGGGTGAGTAGTACCCCTGTTTGGGACAGAGCC	3042
Qy	2853	TAGTGGTAAATCCAGAGAAATGATCCTATCCAAAGCTGATTCACAAACCCACGCTCACC	2912
Db	3043	TAGTGGTAAATCCAAAGAGAAATGATCCTATCCAAAGCTGATTCACAAACCCACGCTCACC	3102
Qy	2913	TCACAGCCGAGGACGAGCATCACTCTGCTGGACGAGCACTTAGGGGCCCTTCCCAAGG	2972
Db	3103	TGACAGCCGAGGACGAGCATCACTCTGCTGGACGAGCACTTAGGGGCCCTTCCCAAGG	3162
Qy	2973	TCTACCTTAGAGCAAAACCGAGTACCTCAGACAGGAAAGTCGGGGCTTTGACCACTACCAT	3032
Db	3163	TCTACCTTAGAGCAAAACCGAGTACCTCAGACAGGAAAGTCGGGGCTTTGACCACTACCAT	3222
Qy	3033	ATCTCGTAGGCCATTTTCTTAGGCATTTGTGAATAGGTAGGTAGTACGTACACACTTTTTCAGA	3092
Db	3223	ATCTCGTAGGCCATTTTCTTAGGCATTTGTGAATAGGTAGGTAGTACGTACACACTTTTTCAGA	3282
Qy	3093	CCAAATTCAAACTGTCTATGACAAAAATCCCGTGGGCTTAGATGGAGATAATTTTTTTTTT	3152
Db	3283	CCAAATTCAAACTGTCTATGACAAAAATCCCGTGGGCTTAGATGGAGATAATTTTTTTTTT	3342
Qy	3153	CTTCTCAGCTTTATGAAGAGAGGGGAACTGTCTAGGATTCAGCTGACACACAGGAAACC	3212
Db	3343	CTTCTCAGCTTTATGAAGAGAGGGGAACTGTCTAGGATTCAGCTGAAACACACAGGAAACC	3402
Qy	3213	TGGCAACATCAGATTTTAAGCTAAGGTTGGGAGGCTTAACGAGCTTACCTCCCTCTTTGTA	3272
Db	3403	TGGCAACATCAGATTTTAAGCTAAGGTTGGGAGGCTTAACGAGCTTACCTCCCTCTTTGTA	3462
Qy	3273	AATCAAGAATTGTTTAAATGGGATTGTCAATCTCTTTAAATAAAGATGAACCTTGGTTTC	3332
Db	3463	AATCAAGAATTGTTTAAATGGGATTGTCAATCTCTTTAAATAAAGATGAACCTTGGTTTC	3522

Search completed: February 27, 2004, 19:06:00  
Job time : 1249 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 27, 2004, 19:06:08 ; Search time 587 Seconds

(without alignments)  
3741.599 Million cell updates/sec

Title: US-09-964-277-21

Perfect score: 517

Sequence: 1 MLPISLQTVFLYFWNRR.....LKGVSQSSFGSMELIEVS 517

Scoring table:

OLIGO  
Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Word size: 1

Total number of hits satisfying chosen parameters: 6745146

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q/cn2\_1/USPTO.spool/US09964277/runat 25022004 085401 2201/app query.fasta.1.711  
-DB=N\_Geneseq 25Jan04 -QFWT=fascap -SUFFIX=olip2n.rng -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL  
-OUTFWT=pto -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09964277 @CGN 1 1 470 @runat 25022004 085401 2201 -NCPU=3  
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREDS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : N\_Geneseq 29Jan04.\*

- 1: Geneseq1980s.\*
- 2: Geneseq1990s.\*
- 3: Geneseq2000s.\*
- 4: Geneseq2001as.\*
- 5: Geneseq2001bs.\*
- 6: Geneseq2002s.\*
- 7: Geneseq2003as.\*
- 8: Geneseq2003bs.\*
- 9: Geneseq2003cs.\*
- 10: Geneseq2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	517	100.0	3332	6	ABK48378
2	488	94.4	2102	7	ADA53105
3	488	94.4	2732	4	AD09492
4	488	94.4	2966	4	AH99685
5	488	94.4	3059	6	AAS15768
6	488	94.4	3104	6	ABN59704
7	488	94.4	3496	6	ABK47596
8	488	94.4	3544	5	AAS14639

9	488	94.4	3544	6	ABK49402	Abk49402 cDNA enco
10	488	94.4	3766	6	ABK14474	Abk14474 Human pro
11	488	94.4	4790	6	ABN83966	Abn83966 Human gen
12	488	94.4	5145	5	ABV20833	Abv20833 Human pro
13	488	94.4	5145	5	ABV21080	Abv21080 Human pro
14	488	94.4	5145	5	ABV26680	Abv26680 Human pro
15	488	94.4	5145	5	ABV20978	Abv20978 Human pro
16	488	94.4	5145	5	ABV21092	Abv21092 Human pro
17	488	94.4	5145	5	ABV21312	Abv21312 Human pro
18	488	94.4	5145	5	ABV21316	Abv21316 Human pro
19	488	94.4	5145	5	ABV26826	Abv26826 Human pro
20	488	94.4	5145	5	ABV27131	Abv27131 Human pro
21	488	94.4	5145	5	ABV26923	Abv26923 Human pro
22	488	94.4	5145	5	ABV27135	Abv27135 Human pro
23	488	94.4	5145	5	ABV28657	Abv28657 Human pro
24	488	94.4	5145	5	ABV22827	Abv22827 Human pro
25	488	94.4	5145	5	ABV26934	Abv26934 Human pro
26	472	91.3	5450	6	ACC60559	Acc60559 Polynucle
27	472	91.3	5450	6	ACC60572	Acc60572 Polynucle
28	455	88.0	2118	4	AAF30479	Aaf30479 Human pro
29	350	67.7	5111	6	ACC60521	Acc60521 Polynucle
30	138	26.7	425	5	ABV10726	Abv10726 Human pro
31	137	26.5	467	5	ABV40998	Abv40998 Human pro
32	137	26.5	467	5	ABV31891	Abv31891 Human pro
33	133	26.5	467	5	ABV40849	Abv40849 Human pro
34	133	25.7	408	5	ABV10907	Abv10907 Human pro
35	132	25.5	438	5	ABV32059	Abv32059 Human pro
36	67	13.0	355	5	ABV01738	Abv01738 Human pro
37	63	12.2	2756	6	ACC60560	Acc60560 Polynucle
38	58	11.2	411	5	ABV01557	Abv01557 Human pro
39	50	9.7	301	4	AAK53879	Aak53879 Murine tr
40	45	8.7	749	4	AAH06539	Aah06539 Human cDN
41	42	8.1	345	5	ABV02076	Abv02076 Human pro
42	42	8.1	346	5	ABV11245	Abv11245 Human pro
43	42	8.1	377	5	ABV41320	Abv41320 Human pro
44	42	8.1	377	5	ABV32391	Abv32391 Human pro
45	36	7.0	419	5	ABV32042	Abv32042 Human pro

ALIGNMENTS

RESULT 1  
ABK48378

ID ABK48378 standard; cDNA; 3332 BP.

XX AC ABK48378;

XX 02-JUL-2002 (first entry)

XX DE cDNA encoding human DSP-16 alternative form protein.

XX KW Human; dual-specificity phosphatase-3; DSP-16; autoimmune disease;  
XX KW mitogen-activated protein kinase; MAP; Duchenne muscular dystrophy;  
XX KW cancer; graft-versus-host disease; allergy; metabolic disease;  
XX KW abnormal cell growth; abnormal cell proliferation; contact inhibition;  
XX KW cell cycle abnormality; anchorage independent cell growth; apoptosis;  
XX KW intercellular adhesion; DSP-16 modulator; chromosome 12p; gene; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX CDS 842..2395

XX FT /\*tag=a

XX FT /product= "Human dual-specificity phosphatase-3 (DSP-16)  
XX FT alternative form protein"

XX WO200226997-A2.

XX PD 04-APR-2002.

XX PF 25-SEP-2001; 2001WO-US030124.

XX PR 26-SEP-2000; 2000US-0235487P.

XX (CEPT-) CEPTTR INC.  
 PA Luche RM, Wei B;  
 XX WPI; 2002-315802/35.  
 XX P-PSDB; AAU79159.  
 DR New DSP-16 polypeptide, useful for identifying modulators of its  
 XX activity, which can be used in the treatment of disorders such as  
 PT Duchenne muscular dystrophy, or cancer.  
 XX Claim 56; Fig 3; 87pp; English.  
 XX The present invention relates to a new polypeptide, DSP-16, having a 665  
 CC amino acid sequence, given in the specification, or a variant having at  
 CC least 50 % identical residues, which retains the ability to  
 CC dephosphorylate an activated mitogen-activated protein (MAP) kinase. The  
 CC invention provides an activated mitogen-activated protein (MAP) kinase. The  
 CC activity, for modulation of a proliferative response in a cell, survival  
 CC of a cell, or differentiation of a cell. The cell displays contact  
 CC inhibition of cell growth or anchorage independent growth and may display  
 CC altered intercellular adhesion. The agent may modulate apoptosis, or the  
 CC cell cycle. The identified modulators can be used to treat Duchenne  
 CC muscular dystrophy, cancer, graft-versus-host disease, autoimmune  
 CC diseases, allergies, metabolic diseases, abnormal cell growth, abnormal  
 CC cell proliferation, and cell cycle abnormalities. The present nucleic  
 CC acid sequence is that of the human dual-specificity phosphatase-3 (DSP-  
 CC 16) gene located on chromosome 12p. This sequence encodes the human DSP-  
 CC 16 alternative form protein of the invention  
 XX  
 XX Sequence 3332 BP; 909 A; 805 C; 823 G; 795 T; 0 U; 0 Other;  
 SQ  
 Alignment Scores:  
 Pred. No.: 0 Length: 3332  
 Score: 517.00 Matches: 517  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0  
 US-09-964-277-21 (1-517) x ABX48378 (1-3332)  
 QY 1 MetLeuProLeuSerLeuGlnThrValPheSerLeuTyrrPheTrpValAsnTrpArgArg 20  
 Db 842 ATGTTGCCCTCTCTCTTCAGACACTGTTTCTCAGCTGATCTTCTGGGTAAACTGGAGAGA 901  
 QY 21 AlaSerThrLeuPheThrCysLeuGlnGluLeuMetGlnGlnAsnGlyLeGlyTyrrVal 40  
 Db 902 GCTTCAACTCTGTTTCACTGCTTTCAGGAGCTGATGTCAGCAATGGGATTTGGTATGTG 961  
 QY 41 LeuAsnAlaSerAsnThrCysProLysProAspPheLeuProGluSerHisPheLeuArg 60  
 Db 962 TTAATAGCCAGCAATACCTGTCCAAAGCCTGACTTATATCCCGAGCTCTCATTTCTCGGT 1021  
 QY 61 ValProValAsnAspSerPheCysGluLysIleLeuProTrpLeuAspLysSerValAsp 80  
 Db 1022 GTCCCTGTGATGACAGCTTTTGTGAGAAATTTTTCCTGCTGGTGGACAAATCACTAGAT 1081  
 QY 81 PheileGluLysAlaLysAlaSerAsnGlyCysValLeuValHisCysLeuAlaGlyIle 100  
 Db 1082 TTCATTGAGAAAGCAAAAGCCTCCAAATGGATGTGTTCTTAGTCACCTGTTTAGCTGGATC 1141  
 QY 101 SerArgSerAlaThrIleAlaIleAlaTyrrIleMetLysArgMetAspMetSerLeuAsp 120  
 Db 1142 TCCCGCTCCGCCACCATCGGTATCGGCTACATCATGAAGAGATGGACATGCTTTAGAT 1201  
 QY 121 GluAlaTyrrArgPheValLysGluLysArgProThrIleSerProAsnPheAsnPheLeu 140  
 Db 1202 GAAGCTTTACAGATTTGTGAAAGAAAAAGAACCTTACTATATCTCCAAACTTCAATTTCTG 1261  
 QY 141 GlyGlnLeuLeuAspTyrrGluLysValIleLysAsnGlnThrGlyAlaSerGlyProLys 160

## RESULT 2

ADA53105  
ID ADA53105 standard; cDNA; 2102 BP.  
XX  
AC ADA53105;  
XX  
DT 20-NOV-2003 (first entry)  
XX  
DE Human coding sequence, SEQ ID 673.  
XX  
KW Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;  
KW Gene Therapy; human; secretory protein; membrane proteins; cancer;  
KW inflammatory disease; osteoporosis; neurological disease; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN EPI293569-A2.  
XX  
PD 19-MAR-2003.  
XX  
PF 21-MAR-2002; 2002EP-00006586.  
XX  
PR 14-SEP-2001; 2001JP-00328381.  
PR 24-JAN-2002; 2002US-0350435P.  
XX  
XX (HELI-) HELIX RES INST.  
PA (REAS-) RES ASSOC BIOTECHNOLOGY.  
XX  
XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;  
PI Yamamoto J, Isono Y, Otsuka K, Nagai K, Irie R, Tamechika I;  
PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;  
XX  
DR WPI: 2003-395539/38.  
DR P-PSDB; ADA54744.  
XX  
XX New polynucleotides encoding full-length polypeptides, e.g. secretory  
PT and/or membrane proteins, useful for developing medicines for diseases in  
PT which the gene is involved, or as target molecules for gene therapy.  
XX  
XX Claim 1; SEQ ID NO 673; 205pp; English.  
XX  
XX The present invention relates to novel human secretory or membrane  
CC proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-  
CC ADA54071). The coding sequences are useful in the gene therapy of  
CC diseases caused by abnormalities of the proteins, e.g. cancer,  
CC inflammatory diseases, osteoporosis or neurological disease.  
XX  
SQ Sequence 2102 BP; 542 A; 546 C; 510 G; 504 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 0 Length: 2102  
Score: 488.00 Matches: 488  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 94.39% Indels: 0  
DB: 7 Gaps: 0

US-09-964-277-21 (1-517) x ADA53105 (1-2102)

QY 30 GluLeuMetGlnGlnAsnGlyLeuGlyTyrValLeuAsnAlaSerAsnThrCysProLys 49  
DB 587 GAGCTGATGCAGCAGATGGGATGGTATGTTGTTAAATGCCAGCATACTGTCCAAAG 646  
QY 50 ProAspPheIleProGluSerHisPheLeuArgValProValAsnAppSerPheCysGlu 69  
DB 647 CCTGACTTTATCCCGAGTCTCAATTCCTGCGTGTGCTGTGTGATGACAGCTTTGTGAG 706  
QY 70 LysIleLeuProTrpLeuAspLysSerValAspPheIleGluLysAlaLysAlaSerAsn 89  
DB 707 AAAATTTTCCCGTGTGGCAAAATCAGTAGATTTCATTGGAAAGCAAAGGCTCCAAAT 766  
QY 90 GlyCysValLeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAla 109

Db 767 GGATGTGTTCTAGTGACATGTTTAGTGGGATCTCCGCTCCGCCACCATCGTATCGCC 826  
QY 110 TyrIleMetLysArgMetAspMetSerLeuAppGluAlaTyrArgPheValLysGluLys 129  
Db 827 TACATCATGAAGAGGATGGACATGTCTTTAGATGAAGCTTACAGATTTTGTGAAGAAAAA 886  
QY 130 ArgProThrIleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGluLysLys 149  
Db 887 AGACCTACTATATCTCAAACCTTCAATTTCTGGGCCAACTCTGGACTATGAGAAGAG 946  
QY 150 IleLysAsnGlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuHisLeuGlu 169  
Db 947 ATTAAGAACCCAGACTGAGCATCAGGGCCAAAGAGCAAACTCAAGCTGTGCACTGGAG 1006  
QY 170 LysProAsnGluProValProAlaValSerGluGlyGlyGlnLysSerGluThrProLeu 189  
Db 1007 AAGCCAAATGAACTGTCCCTGTCTCAGAGGGTGGACAGAAAGCCGAGCCCTTC 1066  
QY 190 SerProCysAlaAspSerAlaThrSerGluAlaAlaGlyGlnArgProValHisPro 209  
Db 1067 AGTCCACCTGTGCCGACTCTGCTACTCAGAGGCGCAGGACAAAGGCCGTGATCC 1126  
QY 210 AlaSerValProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuVal 229  
Db 1127 GCCAGCGTGCACGCGTCCCGCAGCGTGCAGCGCTCGCTGTAGAGACAGCCGCTGTA 1186  
QY 230 GlnAlaLeuSerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnLysLeuLys 249  
Db 1187 CAGCGCTCAGTGGCTGCACCTGTCCGAGACAGGCTGGAGACAGCAATTAAGCTCAAG 1246  
QY 250 ArgSerPheSerLeuAspIleLysSerValSerTyrSerAlaSerMetAlaAlaSerLeu 269  
Db 1247 CGTTCTCTCTCTGGATATCAAACTCAGTTTTCATATTCAGCCAGCATGCGCATCTCTTA 1306  
QY 270 HisGlyPheSerSerSerGluAspAlaLeuGluTyrTyrLysProSerThrThrLeuAsp 289  
Db 1307 CATGGCTTCTCTCATCAGAAAGATGCTTGGATATATACAACTTCCACTACTCTGGAT 1366  
QY 290 GlyThrAsnLysLeuCysGlnPheSerProValGlnGluLeuSerGluGlnThrProGlu 309  
Db 1367 GGGACCAACAAGCTATGCGCAGTTCTCCCTGTTTTCAGAACTATCGGAGCAGACTCCCGAA 1426  
QY 310 ThrSerProAspLysGluGluAlaSerIleProLysLysLeuGlnThrAlaArgProSer 329  
Db 1427 ACCAGTCTCATAGGAGGAAGCCAGCATCCCCAAGAGCTGCGACTGCCAGGCCCTTCA 1486  
QY 330 AspSerGlnSerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArg 349  
Db 1487 GACAGCCAGAGCAAGCGATTGCTTGGTTCAGAACCCAGCAGCAGTGGCACCGCCAGAGG 1546  
QY 350 SerLeuLeuSerProLeuHisArgSerGlySerValGluAspAsnTyrHisThrSerPhe 369  
Db 1547 TCCCTTTTATCTCCACTGCTATCGAAGTGGGAGCGTGAGACAAATACCACACACACTTC 1606  
QY 370 LeuPheGlyLeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLeuLys 389  
Db 1607 CTTTTCGGCTTTCACACGCGCAGCAGCACTCACGAGTCTGCTGGCTGGGCTTAAG 1666  
QY 390 GlyTrpHisSerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTrp 409  
Db 1667 GCGTGGCACTCGGATATCTTGGCCCCCAGACCTCTACCCCTTCCCTGCCACGACGCTGG 1726  
QY 410 TyrPheAlaThrGluSerSerHisPheTyrSerAlaSerAlaIleTyrGlyGlySerAla 429  
Db 1727 TATTTCGCACAGAGTCTCTACACTTCTACTCTGCCTCAGCCATCTACGGAGGACGATGCC 1786  
QY 430 SerTyrSerAlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerVal 449  
Db 1787 AGTTACTCTGCTACAGCTGCAGCAGCTGCCACTTGGAGACCAAGTCTATTCTCTGTG 1846  
QY 450 ArgArgArgGlnLysProSerAspArgAlaAspSerArgSerTrpHisGluLeuSer 469

Db 1847 CCAGCGGCGCAAGCCAAAGTCACAGAGCTGACTCGCGGGAGCTGGCATGAGAGAGC 1906  
Qy 470 ProPheGluLysGlnPheLysArgSerCysGlnMetGluPheGlyGluSerIleMet 489  
Db 1907 CCCTTTGAAAGCAGCTTTAAACGCGAGAGCTGCCAAATGGAATTTGGAGAGACATCATG 1966  
Qy 490 SerGluAsnArgSerArgGluGluLysValGlySerGlnSerSerPheSerGly 509  
Db 1967 TCAGAGACACAGCTCAGCGGAGAGCTGGGAAATGGGAGTCAGTCTAGCTTTTCGGGC 2026  
Qy 510 SerMetGluIleGluValSer 517  
Db 2027 AGCATGGAATCATTTAGGCTCTCC 2050  
RESULT 3  
ID AAD09492 standard; DNA; 2732 BP.  
XX  
AC AAD09492;  
XX  
DT 10-SEP-2001 (first entry)  
XX  
DE Human SGP002 phosphatase polypeptide encoding DNA.  
XX  
KW Human; SGP002 phosphatase polypeptide, phosphatase-related disease;  
KW immune-related disorder; ocular disease; organ transplant rejection;  
KW infection; diabetes; pain; sexual dysfunction; Alzheimer's disease;  
KW metabolic disorder; haematopoietic cancer; mood disorder; cardiac;  
KW Parkinson's disease; multiple sclerosis; amyotrophic lateral sclerosis;  
KW cardiovascular disease; brain; neuronal-associated disease; dyskinesia;  
KW attention disorder; cognition disorder; psychotic disorder; cytostatic;  
KW neurological disorder; viricide; nootropic; cerebroprotective; therapy;  
KW neuroprotective; antibacterial; vulnary; tranquiliser; antidiabetic;  
KW hypotensive; immunosuppressive; antipsoriatic; analgesic; hypertensive;  
KW antifungal; dual specificity phosphatase; DSP; MAP kinase phosphatase;  
KW MKP; migraine; chromosome 12p11.1-p12.1; ds.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 538..2535  
FT /tag= a  
FT /product= "Human SGP002 phosphatase polypeptide"  
XX  
PN WO200146394-A2.  
XX  
PD 28-JUN-2001.  
XX  
PF 21-DEC-2000; 2000WO-US034736.  
XX  
PR 21-DEC-1999; 99US-0173255P.  
PR 28-DEC-1999; 99US-0175766P.  
PR 25-JAN-2000; 2000US-0178078P.  
PR 31-JAN-2000; 2000US-0179301P.  
XX  
PA (SUGEN-) SUGEN INC.  
XX  
PI Plowman GD, Martinez R, Whyte D, Manning G, Sudareanam S;  
PI Hill RJ, Flanagan P;  
XX  
DR WPI; 2001-418058/44.  
DR P-PSDB; AAE04834.  
XX  
PT Novel phosphatase polypeptide useful for treating cancers, immune-related  
PT diseases and disorders, cardiovascular disease, brain or neuronal-  
PT associated diseases and metabolic disorders.  
XX  
PS Claim 29; Fig 1; 186pp; English.  
XX  
CC The present invention relates to phosphatase polypeptides, nucleotide  
CC sequences encoding them, as well as various products and methods useful  
CC for the diagnosis and treatment of various phosphatase-related diseases  
CC and conditions. Substance that modulates the activity of phosphatase

CC polypeptide is used to treat immune-related diseases and disorders,  
CC cardiovascular disease, brain or neuronal-associated diseases and  
CC haematopoietic origin, diseases of central and peripheral nervous system,  
CC Alzheimer's disease, Parkinson's disease, multiple sclerosis, amyotrophic  
CC lateral sclerosis, viral infections, infections caused by prions,  
CC bacteria and fungi, ocular diseases, diabetes, migraines, pain, sexual  
CC dysfunction, mood disorders, attention disorders, neurological disorders,  
CC dyskinesia and organ transplant rejection. The present sequence is a DNA  
CC encoding human SGP002 phosphatase polypeptide. This sequence is  
CC classified as dual specificity phosphatase (DSP) and MAP kinase  
CC phosphatase (MKP). SGP002 gene maps to chromosomal position 12p11.1-p12.1  
XX  
XX Sequence 2732 BP; 710 A; 684 C; 686 G; 652 T; 0 U; 0 Other;  
Alignment Scores:  
Pred. No.: 0 Length: 2732  
Score: 488.00 Matches: 488  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 94.39% Indels: 0  
DB: 4 Gaps: 0  
US-09-964-277-21 (1-517) x AAD09492 (1-2732)  
Qy 30 GluLeuMetGlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerAsnThrCysProLys 49  
Db 1069 GAGCTGATGCGACGAGAATGGGATTGGTTATGTGTTAAATCCGCAATACCTCTCCAAAG 1128  
Qy 50 ProAspPheIleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGlu 69  
Db 1129 CTTGACTTTATCCCGAGTCTCATTTCTCGGTGGCTGCTGATGACAGCTTTTGAG 1188  
Qy 70 LysIleLeuProTyrLeuAspLysSerValAspPheIleGlyLysAlaLysAlaSerAsn 89  
Db 1189 AAAATTTTGGCGTGGTGGACAAATCAGTAGATTTCATTGAGAAAGCAAAGCTCCAT 1248  
Qy 90 GlyCysValLeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAla 109  
Db 1249 GGATGTGTCTAGTGCACCTGTTTAGCTGGGATCTCCCGCTCCGCCACCATCGTATCGCC 1308  
Qy 110 TyrIleMetLysArgMetAspMetSerLeuAspGluAlaTyrArgPheValLysGluLys 129  
Db 1309 TACATCATGAGAGGATGGACATGCTCTTTAGTAGAAGCTTACAGATTGTGAAAGAAAA 1368  
Qy 130 ArgProThrIleSerProAsnPheAsnPheLeuGlyClnLeuLeuAspTyrGluLysLys 149  
Db 1369 AGACCTACTATATCTCCAAACTTCAATTTTCTGGGCCAACTCTCTGGACTATGAGAAAG 1428  
Qy 150 IleLysAsnGlnThrGlyValAserGlyProLysSerLysLysLeuLysLeuHisLeuGlu 169  
Db 1429 ATTAGAACACAGACTGGAGCATCAGGCCCAAGAGCAAACTCAAGCTGCTGCACCTGGAG 1488  
Qy 170 LysProAsnGluProValProAlaValSerGluClyGlyGlnLysSerGluThrProLeu 199  
Db 1489 AAGCCAAATGAACCTGCTCCCTGCTGTCTCAGAGGCTGGACAGAAAGCGAGAGCCCTC 1548  
Qy 190 SerProProCysAlaAspSerAlaThrSerGluAlaAlaGlyGlnArgProValHisPro 209  
Db 1549 AGTCACCTGTGCGGACTCTGCTACTCAGAGCAGCAGGACAAAGCCCGCTGCATCCC 1608  
Qy 210 AlaSerValProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuVal 229  
Db 1609 GCCAGCGTGGCCAGCGTCCCGAGCGTGCAGCGCTGCTGTTTAGAGGACAGCCCGCTCGTA 1668  
Qy 230 GlnAlaLeuSerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnLysLeuLys 249  
Db 1669 CAGCGCGCTCAGTGGGCTGCACCTGTCCGACAGAGCTGGAGACACATATAGCTCAAG 1728  
Qy 250 ArgSerPheSerLeuAspIleLysSerValSerTyrSerAlaSerMetAlaAlaSerLeu 269  
Db 1729 CGTTCTTCTCTCTGATATCAATCATGTTTCATATTACAGCAGCATGGCAGCATCCTTA 1788

QY 270 HisGlyPheSerSerSerGluAspAlaLeuGluTyrTyrIysProSerThrThrLeuAsp 289  
Db 1789 CATGGCTTCTCTCATCAGAGATGCTTGGATACTACTAACCTTCCACTACTCTGGAT 1848  
QY 290 GlyThrAsnLysLeuCyseGlnPheSerProValGlnGluLeuSerGlnThrProGlu 309  
Db 1849 GGGACCAACAGCTATGCGAGTTCTCCCTGTTTCAGGACTATCCGAGCAGATCCCGAA 1908  
QY 310 ThrSerProAspLysGluGluAlaSerIleProLysLysLeuGlnThrAlaArgProSer 329  
Db 1909 ACCAGTCTCTATAGAGGAAGGAGCAGATCCCCAGAGAGCTGCAGACTGCCAGCCCTTCA 1968  
QY 330 AspSerGlnSerLysArgLeuHisSerValArgThrSerSerGlyThrAlaGlnArg 349  
Db 1969 GACACCCAGAGCAAGGATTCATTCGTGTCAGAACCCAGCAGCAGTGCAGACTGCCAGCCCTTCA 1968  
QY 350 SerLeuLeuSerProLeuHisArgSerGlySerValGluAspAsnTyrHisThrSerPhe 369  
Db 2029 TCCCTTTTATCTCCACTGCATCGAAGTGGAGCGTGGAGGACATATACACACAGCTTC 2088  
QY 370 LeuPheGlyLeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLeuLys 389  
Db 2089 CTTTTCGGCCCTTTCACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2148  
QY 390 GlyTyrHisSerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTrp 409  
Db 2149 GGCCTGGCACTCGGATATCTTGGCCCCCAGACCTTACCCCTTCCCTGACACAGCAGCTG 2208  
QY 410 TyrPheAlaThrGluSerSerHisPheTyrSerAlaSerAlaIleTyrGlySerAla 429  
Db 2209 TATTTTGGCACAGAGTCTCTCACTTCTACTCTGCTCAGCCATCTACGGAGCAGCTGCC 2268  
QY 430 SerTyrSerAlaTyrSerCyseGlnLeuProThrCyseGlyAspGlnValTyrSerVal 449  
Db 2269 AGTTACTTGTCTCAGTGCAGCCAGCTGCCCTTCTGCGAGACCAAGTCTATTCGTG 2328  
QY 450 ArgArgArgGlnLysProSerAspArgAlaAspSerArgSerTrpHisGluSer 469  
Db 2329 CGCAGCGCGCAGAACCCAAAGTGACAGAGCTGACTCGCGCGGAGCTGGCATGAGAGAGC 2388  
QY 470 ProPheGluLysGlnPheLysArgSerCyseGlnMetGluPheGlyGluSerIleMet 489  
Db 2389 CCCTTTGAAAGACATTTAAACGCGAGAGCTGCCAATGGAAATTTGAGAGAGCATCATG 2448  
QY 490 SerGluAsnArgSerArgGluGluLeuGlyLysValGlySerGlnSerSerPheSerGly 509  
Db 2449 TCAGAGACAGCTCAGCGGAGAGCTGGGAAAGTGGGCACTAGTCTAGCTTTTCGGC 2508  
QY 510 SerMetGluIleIleGluValSer 517  
Db 2509 AGCATGGAAATCATTTGAGGTCTCC 2532

RESULT 4  
AAH99685  
ID AAH99685 standard; cDNA; 2966 BP.  
XX  
AC AAH99685;  
XX  
DT 16-OCT-2001 (first entry)  
XX  
DE Human protein encoding cDNA sequence SEQ ID NO:520.  
XX Human; cancer; ulcer; HIV infection; human immunodeficiency virus;  
XX antiinflammatory; antirheumatic; antiallergic; immunosuppressive;  
XX antibacterial; endocrine; cardiant; central nervous system; virucide;  
XX anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;  
XX antiaggregant; haemostatic; vulnery; antitumor; osteopathic; eczema;  
XX dermatological; antiallergic; antiasthmatic; antidiabetic; cytostatic;  
XX neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;  
XX immunostimulant; gene therapy; antineoplastic; vaccine; inflammation;  
XX antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;  
XX cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;

KW genetic disease; haematopoietic disorder; platelet disorder; asthma;  
KW thrombocytopenia; osteoporosis; severe combined immunodeficiency;  
KW allergic rhinitis; diabetes; multiple sclerosis; depression;  
KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;  
KW neurological disorder; ss.  
XX Homo sapiens.  
XX WO200153455-A2.  
PN 26-JUL-2001.  
XX  
PF 22-DEC-2000; 2000WO-US035017.  
XX  
PR 23-DEC-1998; 89US-00471275.  
PR 21-JAN-2000; 2000US-00488725.  
PR 25-APR-2000; 2000US-00552317.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
XX Tang YT, Liu C, Drmanac RT;  
XX  
XX WPI; 2001-457603/49.  
DR P-PSDB; AAM25744.  
XX  
PT Isolated human polynucleotides encoding polypeptides, useful for the  
PT treatment and diagnosis of e.g. cancer, ulcers and HIV infection.  
XX  
XX Claim 1; Page 578; 1217pp; English.  
XX  
XX AAH99166 to AAH99904 encode the human proteins given in AAM25225 to  
CC AAM25963. The proteins can have activities based on the tissues and cells  
CC they are expressed in, such as: antinflammatory; antirheumatic;  
CC antiallergic; immunosuppressive; antibacterial; endocrine; cardiant;  
CC central nervous system; virucide; anti-HIV; fungicide; antimutagen;  
CC cardiovascular; antianaemic; antiaggregant; haemostatic; vulnery;  
CC antitumor; osteopathic; dermatological; antiallergic; antiasthmatic;  
CC antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;  
CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides  
CC encoding them can be used in gene therapy, antineoplastic therapy and vaccine  
CC production. The proteins and polynucleotides are useful for screening for  
CC agonists or antagonists of a protein and for the treatment and diagnosis  
CC of disorders associated with the activity of a protein e.g. inflammation,  
CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,  
CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal  
CC infections, autoimmunity, genetic diseases, haematopoietic disorders,  
CC anaemia, platelet disorders, thrombocytopenia, wounds, burns, ulcers,  
CC osteoporosis, severe combined immunodeficiency, eczema, allergic  
CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,  
CC Alzheimer's disease, Parkinson's disease, neurodegenerative and  
CC neurological disorders  
XX  
SQ Sequence 2966 BP; 809 A; 735 C; 688 G; 733 T; 0 U; 1 Other;

Alignment Scores:  
Pred. No.: 0 Length: 2966  
Score: 488.00 Matches: 488  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 94.39% Indels: 0  
DB: 4 Gaps: 0

US-09-964-277-21 (1-517) x AAH99685 (1-2966)

QY 30 GluLeuMetGlnGlnAsnGlyLeuGlyTyrValLeuAsnAlaSerAsnThrCysProLys 49  
Db 554 GAGCTGATGACGAGAAATGGGATGGTATGTGTAAATCCAGCAATACCTGTCCAAAG 613  
QY 50 ProAspPheIleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGlu 69  
Db 614 CCGTACTTATCCCGAGTCTCATTTCTCGGTGCTGCTGATGACACTTTTGAG 673  
QY 70 LysIleLeuProTrpLeuAspLysSerValAspPheIleGluLysAlaLysAlaSerAsn 89



```
Db 674 AAAAAATTTTCCCGTGGTGGACAAAATCAGTAGATTTCATTGGAGAAGCAAAAGCCCTCCAAT 733
Qy 90 GlyCysValLeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAla 109
Db 734 GGATGTGTTCTAGTGCACATGTTAGCTGGATCTCCCGCTCCGCCACCATCGCTATCGCC 793
Qy 110 TyrIleMetLysArgMetAspMetSerLeuAspGluAlaTyrArgPheValLysGluLys 129
Db 794 TAGATCATGAAGAGGATGACATGCTTTTAGATGAAGCTTACAGATTGTGAAGAAAAA 853
Qy 130 ArgProThrIleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGluLysLys 149
Db 854 AGACCTACTATATCTCAAACTTCATTTCTGGGCCAACTCCTGGACTATGAGAAGAAG 913
Qy 150 IleLysAsnGlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuLeuHisLeuGlu 169
Db 914 ATTAAAGAACCCAGACTGGAGCATCAGGGCCAAAGAGCAAACTCAAGCTGCTCACCTGGAG 973
Qy 170 LysProAsnGluProValProAlaValSerGluGlyGlyGlnLysSerGluThrProLeu 189
Db 974 AAGCCAAATGAACCTGCTCCTGCTGCTCAGAGGGTGGACAGAAAGCGAGCGCCCTC 1033
Qy 190 SerProProCysAlaAspSerAlaThrSerGluAlaAlaGlyGlnArgProValHisPro 209
Db 1034 AGTCCACCCCTGTCCGACTCTGCTACCTCAGAGGCGAGCAGCAAAAGCCCGCTGCATCC 1093
Qy 210 AlaSerValProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuVal 229
Db 1094 GCCAGCGTGGCCAGCGTGCAGCGTGCAGCGTGCCTGCTTTAGAGCAGAGCCGCTGGTA 1153
Qy 230 GlnAlaLeuSerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnLysLeuLys 249
Db 1154 CAGCGCGCTCAGTGGCTGACCTGTCCGAGACAGCGCTGGAAGACAGCAATAAGCTCAAG 1213
Qy 250 ArgSerPheSerLeuAspLysSerValSerTyrSerAlaSerMetAlaAlaSerLeu 269
Db 1214 CGTTCCTTCTCTCGATATCAAAATCAGTTTCATATTCAGCCAGCATGGCAGCATCCTTA 1273
Qy 270 HisGlyPheSerSerGluAspAlaLeuGluTyrTyrLysProSerThrThrLeuAsp 289
Db 1274 CATGGCTTCTCTCATCAGAGATGCTTTGGATACTACAAACCTTCCACTACTCTGGAT 1333
Qy 290 GlyThrAsnLysLeuCysGlnPheSerProValGlnGlnLeuSerGluGlnThrProGlu 309
Db 1334 GGGACCAACCAAGCTATGCCAGTTCTCCCTGTTCAGGAACTATCGGAGCAGACTCCCGAA 1393
Qy 310 ThrSerProAspLysGluGluAlaSerIleProLysLysLeuGlnThrAlaArgProSer 329
Db 1394 ACCAGTCTCATAGGAGGAGGAGCAGCATCCCCAGAGCTGCAGACCGCCAGGCTTCA 1453
Qy 330 AspSerGlnSerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArg 349
Db 1454 GACAGCCAGAGCAAGCATTCATTCGCTCAGAACCCAGCAGCAGTGGCACCAGCCAGGAG 1513
Qy 350 SerLeuLeuSerProLeuHisArgSerGlySerValGluAspAsnTyrHisThrSerPhe 369
Db 1514 TCCCTTTTATCTCCACTGCATCGAATGGAGGCGGTGGAGGCAATTTACACACCGAGCTTC 1573
Qy 370 LeuPheGlyLeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLeuLys 389
Db 1574 CTTTTTCGGCTTTCACACCGCAGCAGCAGCTCAACGAGTCTGCTGGCTGGCCTTAAG 1633
Qy 390 GlyThrHisSerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerThr 409
Db 1634 GGCTGGCACTCGGATATCTTGGCCCCCAGACCTCTACCCCTTCCCTGACACGAGCTGG 1693
Qy 410 TyrPheAlaThrGluSerSerHisPheTyrSerAlaSerAlaIleTyrGlySerAla 429
Db 1694 TATTTTCCACAGAGTCTCTCAGCTTCTACTCTGCTCTAGCCATCTACGAGGAGCTGCC 1753
Qy 430 SerTyrSerAlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerVal 449
```

1754 AGTTACTCTGCTACAGCTGCAGCCAGCTGCCACTTTCGGAGACCAAGTCTATTCTGTG 1813

450 ArgArgArgGlnLysProSerAspArgAlaAspSerArgSerTrpHisGluGluSer 469

1814 CGCAGCGCGCAGAAAGCCAAAGTGACAGAGCTGACTCGCGCGGAGCTGGCATGAAGAGC 1873

470 ProPheGluLysGlnPheLysArgSerCysGlnMetGluPheGlyGluSerIleMet 489

1874 CCCTTTGAAAGCAGTTTAAACGCGAAGCTGCCAAATGGAAATTTGGAGAGACATCATG 1933

490 SerGluAsnArgSerArgGluGluLeuGlyLysValGlySerGlnSerSerPheSerGly 509

1934 TCAGAGAAACAGCTCACGGGAAAGAGCTGGGAAAGTGGCGAGTCAGTCTAGCTTTTCGGGC 1993

510 SerMetGluIleIleGluValSer 517

1994 AGCATGGAATCATTTAGGTCTCC 2017

RESULT 5

AAS15768

ID AAS15768 standard; cDNA; 3059 BP.

XX

AC AAS15768;

DT 18-JUN-2002 (first entry)

XX

DE cDNA sequence of human (dual specificity phosphatase) DUSP-10.

XX

KW Human; dual specificity phosphatase; DUSP-10; cancer; epilepsy; stroke; neuronal degeneration syndrome; Alzheimer's disease; depression; schizophrenia; asthma; immune disorder; inflammatory process; arthritis; osteoporosis; diabetes; gene therapy; chromosome 12; ss.

XX

OS Homo sapiens.

XX

PH Key

FT CDS

FT 127..2124

FT /\*tag= a

FT /product= "DUSP-10 protein"

PN WO200177340-A1.

XX

PD 18-OCT-2001.

XX

PF 06-APR-2001; 2001WO-EP003966.

XX

PR 10-APR-2000; 2000EP-00107143.

XX

PA (MERE ) MERCK PATENT GMBH.

XX

PI Duecker K;

XX

DR WPI; 2002-010917/01.

DR P-PSDB; AAU09946.

XX

PT Novel dual specificity phosphatase polypeptides useful for treating cancers, epilepsy, stroke, neuronal degeneration syndromes, Alzheimer's disease, depression, schizophrenia, asthma and immune disorders.

PS Claim 5; Page 34-37; 43pp; English.

XX

CC The present invention relates to a new isolated dual specificity phosphatase (DUSP10) polypeptide, comprising a 665 residue amino acid sequence that is fully defined in the specification. The invention also provides a sequence encoded by a 3059 nucleotide sequence fully defined in the specification, and a sequence having at least 95 % identity to the polypeptide, or fragments or variants of DUSP-10. The invention is useful for treating cancer e.g. leukaemia, colon carcinoma, lung cancer, prostate cancer, metastasis of tumour cells, neo-angiogenesis, epilepsy, stroke, neuronal degeneration syndromes, Alzheimer's disease, depression, schizophrenia, cardiac myopathies, asthma, immune disorders, inflammatory processes e.g. arthritis, bowel disease, type 1 diabetes, osteoporosis, diabetes and diabetes associated diseases. The molecules of

CC the invention are also useful as vaccines for inducing immunological  
CC response in a mammal, in disease diagnosis and in assays for screening  
CC agonistic or antagonistic compounds. Other uses of the invention include  
CC identifying membrane bound or soluble receptors, as a diagnostic reagent,  
CC in chromosome localisation studies, and as a valuable tool in tissue  
CC expression studies. The present sequence represents cDNA of the human  
CC dual specificity phosphatase, DUSP-10. DUSP-10 is located on chromosome  
CC 12

XX Sequence 3059 BP; 831 A; 761 C; 709 G; 758 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 0 Length: 3059  
Score: 488.00 Matches: 488  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 94.39% Indels: 0  
DB: 6 Gaps: 0

US-09-964-277-21 (1-517) x AAS15768 (1-3059)

Qy 30 GluLeuMetGlnGlnAenGlyIleGlyTyrValLeuAsnAlaSerAsnThrCysProLys 49  
Db 658 GAGCTGATGAGCAGAGATGGATGGTATGTTAAATGCCAGCAATACCTGTCRAAG 717  
Qy 50 ProAspPheileProGluSerHisPheLeuArgValProValAsnAspSerPheCysGlu 69  
Db 718 CTTGACATTTATCCCGAGTCTCATTTCTCGGTGCTGTGAATGACAGCTTTGTGAG 777  
Qy 70 LysileLeuProThrLeuAspLysSerValAspPheIleGluLysAlaLysAlaSerAsn 89  
Db 778 AAAATTTTGGCGTGGTGGCAATCATAGTAGATTTTCAATTGAGAAAGCAAAAGCCCTCAAT 837  
Qy 90 GlyCysValLeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAla 109  
Db 838 GGAATGTTTCTAGTGCACTGTTTACCTGGGATCTCCGCTCCGCCCACTTCGTATCGCC 897  
Qy 110 TyrIleMetLysArgMetAspMetSerLeuAspGluAlaTyrArgPheValLysGluLys 129  
Db 898 TACATCATGAAGAGATGGACATGCTCTTAGATGAAGCTTACAGATTTGTGAAGAAAA 957  
Qy 130 ArgProThrIleSerProAsnPheAsnPheLeuGlyGlnLeuAspTyrGluLysLys 149  
Db 958 AGACCTACTATCTCCAAATCTCAATTTCTGGGCCAATCTCTGGACTATGAGAAGAG 1017  
Qy 150 IleLysAsnGlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuHisLeuGlu 169  
Db 1018 ATTAGAACCACTGGAGCATCAGGCCCAAGAGCAACTCAAGCTGTCACCTGGAG 1077  
Qy 170 LysProAsnGluProValProAlaValSerGluGlyGlnLysSerGluThrProLeu 189  
Db 1078 AAGCCAAATGAACCTGTCCCTGTCTCAGAGGGTGGACAGAAAAAGCGAGACGCCCTC 1137  
Qy 190 SerProCysAlaAspSerAlaThrSerGluAlaAlaGlyGlnArgProValHisPro 209  
Db 1138 AGTCCACCTGTGGCACTGTGCTACTCTCAGAGCAGCAGCAAAAGGCCGTGCACTCC 1197  
Qy 210 AlaSerValProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuVal 229  
Db 1198 GCCAGCGTCCCGAGCGTCCCGAGCGTGGCGTGTGTAGAGCAGACGCCCGCTGGTA 1257  
Qy 230 GlnAlaLeuSerGlyLeuHisLeuSerAlaAspArgLeuAspSerAsnLysLeuLys 249  
Db 1258 CAGGCGCTCAGTGGGCTGCACCTGTCGCGAGACAGGCTGGAAGACAGCAATAGCTCAAG 1317  
Qy 250 ArgSerPheSerLeuAspIleLysSerValSerTyrSerAlaSerMetAlaAlaSerLeu 269  
Db 1318 CGTTCCTTCTCTGGATATCAAAATCAGTTTCATATTATTCAGCCAGCATGGCAGCATCTTA 1377  
Qy 270 HisGlyPheSerSerGluAspAlaLeuGluTyrTyrLysProSerThrThrLeuAsp 289  
Db 1378 CATGCTCTCTCATCAGAAAGATGCTTTGGAATACTACAAACCTTCCACTACTCTGGAT 1437

Qy 290 GlyThrAsnLysLeuCysGlnPheSerProValGlnGluLeuSerGluGlnThrProGlu 309  
Db 1438 GGACCAAAAGCTATGCGCAGTTCTCCCTGTTCAGGAACATTCGAGCAGACTCCCCAA 1497  
Qy 310 ThrSerProAspLysGluAlaSerIleProLysLeuLysLeuGlnThrAlaArgProSer 329  
Db 1498 ACCAGTCTCTGATAAGGAGGAGCAGCATCCCAAGAGCTCAGACCGCAGGCTTCA 1557  
Qy 330 AspSerGlnSerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArg 349  
Db 1558 GACAGCCAGCAGCAAGCGATTGCTTCGTCAGAACAGCAGCAGTGGCACCCCGCAGAG 1617  
Qy 350 SerLeuLeuSerProLeuHisArgSerGlySerValGluAspAsnTyrHisThrSerPhe 369  
Db 1618 TCCCTTTTATCTCATCTGCATCGAGTGGAGCGTGGAGCAATTACCACACAGCTTC 1677  
Qy 370 LeuPheGlyLeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLeuLys 389  
Db 1678 CTTTTCGGCCTTTCCACAGCCAGCAGCACCTCAGAAAGTCTGCTGGCCTGGGCTTAAG 1737  
Qy 390 GlyTrpHisSerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTrp 409  
Db 1738 GGTGGCACTCGGATATCTTGGCCCCCAGACCTTACCCCTTCCCTGACACAGCAGCTGG 1797  
Qy 410 TyrPheAlaThrGluSerSerHisPheTyrSerAlaSerAlaIleTyrGlyGlySerAla 429  
Db 1798 TATTTTGGCCACAGAGTCTCACAATCTTACTCTGCTCAGCCATCTACGGAGCGAGTGCC 1857  
Qy 430 SerTyrSerAlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerVal 449  
Db 1858 AGTTACTCTGCTTACAGCTGACCGCAGCTGCCCACTTGGCGAGACAAGTCTATTCTGTG 1917  
Qy 450 ArgArgGlnLysProSerAspArgAlaAspSerArgArgSerTrpHisGluGluSer 469  
Db 1918 CGCAGCGCGCAGAGCAAGTGCAGAGCTGACTCGCGCGAGCTGGCATGAAGAGAGC 1977  
Qy 470 ProPheGluLysGlnPheLysArgArgSerCysGlnMetGluPheGlyGluSerIleMet 489  
Db 1978 CCCTTTGAAAGCAGTTTAAACGACAGAGCTGCCAAATGGAAATTTGGAGAGAGCATCATG 2037  
Qy 490 SerGluAsnArgSerArgGluLeuLeuGlyLysValGlySerGlnSerSerPheSerGly 509  
Db 2038 TCAGAGNACAGCTCAGCGGAGAGAGCTGGGGAAGTGGGCAAGTGGCAGTCTTTTCGGGC 2097  
Qy 510 SerMetGluIleLeuGluValSer 517  
Db 2098 AGCATGGAATCATTCAGTCTCC 2121  
RESULT 6  
ABNS9704  
ID ABNS9704 standard; cDNA; 3104 BP.  
AC ABNS9704;  
XX  
XX 28-JUN-2002 (first entry)  
XX  
XX Novel human coding sequence SEQ ID NO: 115.  
DE  
XX Human; antianaemic; vulnery; antiinflammatory; immunomodulator;  
XX antiinfertility; cerebroprotective; cytostatic; rheumatic; gene therapy;  
KW neuroprotective; antiparkinsonian; protein therapy; EST;  
KW expressed sequence tag; gene; ss.  
XX Homo sapiens.  
XX  
XX WO200222660-A2.  
PN  
XX  
XX 21-MAR-2002.  
PD  
XX  
XX 10-SEP-2001; 2001WO-US026015.  
PF  
XX  
XX 11-SEP-2000; 2000US-00659671.  
PR  
XX



XX Human; dual-specificity phosphatase-3; DSP-16; autoimmune disease;  
KW mitogen-activated protein kinase; MAP; Duchenne muscular dystrophy;  
KW cancer; graft-versus-host disease; allergy; metabolic disease;  
KW abnormal cell growth; abnormal cell proliferation; contact inhibition;  
KW cell cycle abnormality; anchorage independent cell growth; apoptosis;  
KW intercellular adhesion; DSP-16 modulator; chromosome 12p; gene; ss.  
XX  
OS Homo sapiens.  
XX  
XX  
FH Key Location/Qualifiers  
FT CDS 562..2559  
FT /tag= a  
FT /product= "Human dual-specificity phosphatase-3 (DSP-16)  
FT protein"  
XX  
PN WO200226997-A2.  
XX  
PD 04-APR-2002.  
XX  
PF 25-SEP-2001; 2001WO-US030124.  
XX  
PR 26-SEP-2000; 2000US-0235487P.  
XX  
XX (CEPT-) CEPTYR INC.  
XX  
XX Luche RM, Wei B;  
PI WPI; 2002-315802/35.  
DR P-PSDB; AAU79156.  
XX  
XX New DSP-16 polypeptide, useful for identifying modulators of its  
PT activity, which can be used in the treatment of disorders such as  
PT Duchenne muscular dystrophy, or cancer.  
XX  
XX Claim 7; Fig 1; 97pp; English.  
XX  
CC The present invention relates to a new polypeptide, DSP-16, having a 665  
CC amino acid sequence, given in the specification, or a variant having at  
CC least 50 % identical residues, which retains the ability to  
CC dephosphorylate an activated mitogen-activated protein (MAP) kinase. The  
CC invention can be used for identifying agents which modulate DSP-16  
CC activity, for modulation of a proliferative response in a cell, survival  
CC of a cell, or differentiation of a cell. The cell displays contact  
CC inhibition of cell growth or anchorage independent growth and may display  
CC altered intercellular adhesion. The agent may modulate apoptosis, or the  
CC cell cycle. The identified modulators can be used to treat Duchenne  
CC muscular dystrophy, cancer, graft-versus-host disease, autoimmune  
CC diseases, allergies, metabolic diseases, abnormal cell growth, abnormal  
CC cell proliferation, and cell cycle abnormalities. The present nucleic  
CC acid sequence is that of the human dual-specificity phosphatase-3 (DSP-  
CC 16) gene located on chromosome 12p. This sequence encodes the human DSP-  
CC 16 protein of the invention  
XX  
SQ Sequence 3496 BP; 939 A; 854 C; 856 G; 847 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 0 Length: 3496  
Score: 488.00 Matches: 488  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 94.39% Indels: 0  
DB: 6 Gaps: 0

US-09-964-277-21 (1-517) x ABK47596 (1-3496)

QY 30 GluLeuMetGlnGlnAsnGlyLeuGlyTyrValLeuAsnAlaSerAsnThrCysProLys 49  
DB 1093 GAGCTGATGCAGCAGATGGGATGGTATGTTGTTAAATGCCAGCAATACCTGTCCTCAAG 1152

QY 50 ProAspPheLeuProGluSerHisPheLeuArgValProValAsnAspSerPheCysGlu 69  
DB 1153 CCTGACATTTATCCCCGAGTCTCATTTCTCCCTCGGTGCTGTGAATGACAGCTTTGTGAG 1212

QY 70 LysIleLeuProThrIleuAspLysSerValAspPheIleGlyLysAlaLysAlaSerAsn 89  
DB 1213 AAAATTTTGGCGTGTGGCAAAATCAGTAGATTTCATTAGAAAGCAAAAGGCTCCAAT 1272

QY 90 GlyCysValLeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAla 109  
DB 1273 GGATGTGTTCTAGTGCACTGTTTAGCTGGGATCTCCCGCTCCGCCACCATCGCTATCGCC 1332

QY 110 TyrIleMetLysArgMetAspMetSerLeuAspGluAlaTyrArgPheValLysGluLys 129  
DB 1333 TACATCATGAAGAGGATGGACATGTCTTTAGATGAAGCTTACAGATTTGTGAAGAAAAA 1392

QY 130 ArgProThrIleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGlyLysLys 149  
DB 1393 AGACCTACTATATCTCCAACTTCAATTTCTGGGCCAACTCTCTGGACTATGAGAAGAG 1452

QY 150 IleLysAsnGlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuHisLeuGlu 169  
DB 1453 ATTAAGAACCAAGATGGAGCATCAGGGCCAAAGAGCAAACTCAAGCTGTGCACCTGGAG 1512

QY 170 LysProAsnGluProValProAlaValSerGluGlyGlyGlnLysSerGluThrProLeu 189  
DB 1513 AAGCCAAATGAACCTGTCCTGCTCTCAGAGGGTGGACAGAAAGCGAGAGCCCTTC 1572

QY 190 SerProProCysAlaAspSerAlaThrSerGluAlaAlaGlyGlnArgProValHisPro 209  
DB 1573 AGTCCACCTGTGCGGACTCTGCTACTCAGAGGCGAGGAGCAAAAGGCCCTGTGCATCCC 1632

QY 210 AlaSerValProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuVal 229  
DB 1633 GCCAGCTGCTCCAGCGTGCACCGTGCAGCGCTGCTGTAGAGGACAGCCGCTGGTA 1692

QY 230 GlnAlaLeuSerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnLysLeuLys 249  
DB 1693 CAGCGCTCAGTGGCTGTCACCTGTCCGCGAGCAGGCTGGAGACAGCAATAAGCTCAAG 1752

QY 250 ArgSerPheSerLeuAspLysSerValSerTyrSerAlaSerMetAlaAlaSerLeu 269  
DB 1753 CGTTCTCTCTCTGGATATCAATCAGTTTCATATTCAGCCAGCATGGCAGCATCTTTA 1812

QY 270 HisGlyPheSerSerGluAspAlaLeuGluTyrTyrLysProSerThrThrLeuAsp 289  
DB 1813 CATGGCTTCTCTCATCAGAGATGTTTGGAAATACTACAAACCTTCCACTACTCTGGAT 1872

QY 290 GlyThrAsnLysLeuCysGlnPheSerProValGlnGluLeuSerGluGlnThrProGlu 309  
DB 1873 GGGACCAACAAGCTATGCCAGTTCTCCCTGTTCCAGAACTATCGAGCAGACTCCCGAA 1932

QY 310 ThrSerProAspLysGluGluAlaSerIleProLysLysLeuGlnThrAlaArgProSer 329  
DB 1933 ACCAGTCTCTGATAGGAGGAAGCCAGCATCCCCAAGAGCTGCAGACCCGAGGCTTCA 1992

QY 330 AspSerGlnSerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArg 349  
DB 1993 GACAGCCAGAGCAAGCGATTGCAATTCGGTCAAGAACCCAGCAGCAGTGGCACCCCGAGG 2052

QY 350 SerLeuLeuSerProLeuHisArgSerGlySerValGluAspAsnTyrHisThrSerPhe 369  
DB 2053 TCCCTTTTATCTCTCACTGCATCAGAGTGGGAGCGTGGAGACAATTACCACACCAAGCTTC 2112

QY 370 LeuPheGlyLeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLeuLys 389  
DB 2113 CTTTTGGGCTTTTCCACCGCCAGCAGCACTTCAAGAGTCTGCTGGCTGGGCTTAAAG 2172

QY 390 GlyThrPheSerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTyr 409  
DB 2173 GGCTGGCACTCGGATATCTTGGCCCCCAGAGCTCTACCCCTTCCCTGACCAAGAGCTGG 2232

QY 410 TyrPheAlaThrGluSerSerHisPheTyrSerAlaSerAlaIleTyrGlyGlySerAla 429  
DB 2233 TATTTTGGCCAGAGTCTCTCACTTCTACTCTGCTCTAGCCATCTACGGAGGCGAGTGC 2292

Qy 430 SerTyrSerAlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerVal 449  
Db 2293 AGTTACTTCGCTACAGCTGCAGCCAGCTGCCCACTTGGGAGACCAAGTCTATCTGTG 2352  
Qy 450 ArgArgArgGlnLysProSerAspArgAlaAspSerArgSerTrpHisGluGluSer 469  
Db 2353 CGCAGGCGGAGAGACCAAGTGCAGAGCTGACTCGCGCGAGCTGGCATGAAGAGAGC 2412  
Qy 470 ProPheGluLysGlnPheLysArgSerCysGlnMetGluPheGlyGluSerIleMet 489  
Db 2413 CCTTTGAAAAGCAGTTTAAAGCGCAGAGCTGCCAAATGGAATTTGGAGAGAGCATCATG 2472  
Qy 490 SerGluAsnArgSerArgGluGluLeuGlyLysValGlySerGlnSerSerPheSerGly 509  
Db 2473 TCAGAAACAGGTACGGAGAGAGTGGGGAAGTGGGAGTCAAGTCTAGCTTTTCGGGC 2532  
Qy 510 SerMetGluIleGluValSer 517  
Db 2533 AGCATGGAATCATTTGAGGTCTCC 2556  
RESULT 8  
AAS14639  
ID AAS14639 standard; cDNA; 3544 BP.  
XX  
AC AAS14639;  
DT 18-DEC-2001 (first entry)  
XX  
DE Human cDNA encoding dual specificity phosphatase 21117.  
XX  
KW Human; ss; dual specificity phosphatase 21117; hepatotropic; cytostatic;  
KW hematopoietic disorder; autoimmune disorder; diabetes mellitus;  
KW rheumatoid arthritis; multiple sclerosis; Crohn's disease;  
KW liver disorder; erythroid associated disorder; haemolytic anaemia;  
KW cellular proliferative; differentiative disorder; leukaemia;  
KW acute myeloid leukaemia; carcinoma; sarcoma; metastatic cancer.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 589..2586  
FT /tag= a  
FT /product= "Dual specificity phosphatase 21117"  
XX  
PN WO200173059-A2.  
XX  
XX 04-OCT-2001.  
XX  
XX 23-MAR-2001; 2001WO-US009477.  
XX  
XX 24-MAR-2000; 2000US-0191858P.  
XX  
XX (MILL-) MILLENNIUM PHARM INC.  
XX  
XX Meyers RA;  
XX  
XX WPI; 2001-611635/70.  
XX P-PSDB; AAU09016.  
XX  
XX New human dual specificity polypeptides and nucleic acids for diagnosis  
XX of disease and treatment of e.g. liver disorders.  
XX  
XX Claim 1; Fig 1; 143pp; English.  
XX  
XX The invention relates to two novel human dual specificity phosphatases  
XX designated 21117 and 38692, the nucleic acids encoding them (including  
XX fragments, allelic variants, their complements or nucleic acids that  
XX hybridize to them) and antibodies raised against the proteins. The  
XX antibody is useful for detecting the presence of the polypeptide, and the  
XX nucleic acid fragments are useful for detecting the presence of the  
XX nucleic acid, for diagnosing disease. Agents (e.g. the antibody and  
XX antisense sequences) are useful for modulating the activity or expression  
XX of the polypeptide/nucleic acid, useful for the treatment of e.g.

CC hematopoietic disorders, autoimmune disorders (e.g. diabetes mellitus,  
CC rheumatoid arthritis, multiple sclerosis, Crohn's disease and many others  
CC listed in the specification) liver disorders, erythroid associated  
CC disorders (e.g. haemolytic anaemia) cellular proliferative or  
CC differentiative disorders, leukaemias (e.g. acute myeloid leukaemia),  
CC carcinomas, sarcomas, metastatic cancers. Agents modulating 21117 and  
CC 38692 are also useful for modulating the proliferation, survival, and  
CC migration or differentiation of a 38692 or 21117-expressing cell. The  
CC polypeptide and nucleic acids are useful for identifying modulating  
CC agents. The present sequence encodes the dual specificity phosphatase  
CC 21117  
XX  
SQ Sequence 3544 BP; 1000 A; 840 C; 822 G; 882 T; 0 U; 0 Other;  
Alignment Scores:  
Pred. No.: 0 Length: 3544  
Score: 488.00 Matches: 488  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 94.39% Indels: 0  
DB: 5 Gaps: 0  
US-09-964-277-21 (1-517) x AAS14639 (1-3544)  
Qy 30 GluLeuMetGlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerAsnThrCysProLys 49  
Db 1120 GAGCTGATGCAGCAGATGGGATTGGTTATGTGTTAAATGCCAGCAATACCTGTCCAAAG 1179  
Qy 50 ProAspPheIleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGlu 69  
Db 1180 CCGTACTTTATCCCGAGTCTCATTTCTCGGTGCTGCTGTAATGACAGCTTTTGGAG 1239  
Qy 70 LysIleLeuProTrpLeuAspLysSerValAspPheIleGluLysAlaLysAlaSerAsn 89  
Db 1240 AAAATTTTCGCTGGTGGACAAATCAGTAGATTTCATTGGAAGCAAAAGCTCCAT 1299  
Qy 90 GlyCysValLeuValHisCysLeuAlaGlyLysSerArgSerAlaThrIleAlaIleAla 109  
Db 1300 GGATGTGTTCTAGTGCACCTGTTTAGCTGGGATCTCCCGCTCCGACCATCGTATCGCC 1359  
Qy 110 TyrIleMetLysArgMetAspMetSerLeuAspGluAlaTyrArgPheValLysGluLys 129  
Db 1360 TACATCATGAAGAGATGGACATGTCCTTTAGTAGAAGCTACAGATTGTGGAAGAAAA 1419  
Qy 130 ArgProThrIleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGluLysLys 149  
Db 1420 AGACCTACTATATCTCCAAACTTCAATTTCTGGCCCAACTCTCTGGACTATGAGAAGAG 1479  
Qy 150 IleLysAsnGlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuLeuGlu 169  
Db 1480 ATTAAAGAACCACTGGAGCATCAGGGCCAAAGAGCAAACTCAAGCTCTGCACCTGGAG 1539  
Qy 170 LysProAsnGluProValProAlaValSerGluGlyGlnLysSerGluThrProLeu 189  
Db 1540 AAGCCAAATGAACCTGTCCCTGCTGTCTCAGAGGTGGACAGAAAAGCGAGAGCCCTC 1599  
Qy 190 SerProProCysAlaAspSerAlaThrSerGluAlaAlaGlyGlnArgProValHisPro 209  
Db 1600 AGTCACACCTGTGGAGCTCTGTACCTCAGAGCAGCAGGCAAAAGCCCGTCATCCC 1659  
Qy 210 AlaSerValProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuVal 229  
Db 1660 GCCAGCGTGGCCCGGTGCCAGCGTGCAGCGCTGCTGTAGAGGACAGCCCGCTGTA 1719  
Qy 230 GlnAlaLeuSerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnLysLeuLys 249  
Db 1720 CAGGCGCTCAGTGGCTGTCACCTCTCCGACAGAGGTGGAAAGACAGCAATAAGCTCAAG 1779  
Qy 250 ArgSerPheSerLeuAspIleLysSerValSerTyrSerAlaSerMetAlaAlaSerLeu 269  
Db 1780 CTTTCTCTCTCTGGATATCAATCAGTTTCATTATTCAGCCAGCATGGCAGCATCTTTA 1839  
Qy 270 HisGlyPheSerSerSerGluAspAlaLeuGluTyrTyrLysProSerThrThrLeuAsp 289

```
Db 1840 CAGGCTTCTCTCATCAGAGATGCTTTGGAACTACAAACCTTCCACTACTCTGGAT 1899
QY GlyThrAsnLysLeuCysGlnPheSerProValGlnLeuSerGluGlnThrProGlu 309
Db 1900 GGGACCAACAAGCTATGCCAGTCTCTCCCTGTTCCAGGAATATCGGAGCAGACTCCCGAA 1959
QY ThrSerProAspLysGluGluAlaSerIleProLysLysLeuGlnThrAlaArgProSer 329
Db 1960 ACCAGTCTGTATAGGAGGAGCAGCATCCCCAAGAAGCTGCAGACCCGACGCGCTTCA 2019
QY AspSerGlnSerLysArgLeuHisSerValArgThrSerSerGlyThrAlaGlnArg 349
Db 2020 GACAGCCAGAGCAAGCGATTGCATTGCTCAGAACCCAGCAGCAGTGCCACCGCCAGAGS 2079
QY SerLeuLeuSerProLeuHisArgSerGlySerValGluAspAsnThrHisThrSerPhe 369
Db 2080 TCCCTTTTATCTCCACTCATCGAAGTGGGAGCGGTGGAGAGCAATACCAACACCGCTTC 2139
QY LeuPheGlyLeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLeuLys 389
Db 2140 CTTTCGGCCCTTCCACAGCCAGCAGCACTCACGAAGTCTGCTGGCTGGGCTTAA 2199
QY GlyTrpHisSerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerTrp 409
Db 2200 GGCTGGCACTCGGATATCTTGGCCCCCAGACCTCTACCCCTTCCCTGACAGCAGCTGG 2259
QY TyrPheAlaThrGluSerSerHisPheTySerAlaSerAlaIleTyGlyGlySerAla 429
Db 2260 TATTTTGCACAGAGTCTCACATCTTCTCTCTCCCTCAGCCATCTACGAGGAGCTGCC 2319
QY SerTySerAlaTySerCysSerGlnLeuProThrCysGlyAspGlnValTySerVal 449
Db 2320 AGTTACTCTGCTCAGCTCAGCCAGCTGCCCTGCTGGAGACCAAGTCTATTCTGTG 2379
QY ArgArgArgGlnLysProSerAspArgAlaAspSerArgArgSerTrpHisGluGluSer 469
Db 2380 CGCAGCGCGGAGAGCCAGCAAGTCACAGAGCTGACTCGCGCGGAGCTGGCATGAAGAGAGC 2439
QY ProPheGluLysGlnPheLysArgSerCysGlnMetGluPheGlyGluSerIleMet 489
Db 2440 CCCTTTGAAGAGCATTAAACGCGAGAGCTGCCAATGGATTTGGAGAGACATCATG 2499
QY SerGluAsnArgSerArgGluGluLeuGlyLysValGlySerGlnSerSerPheSerGly 509
Db 2500 TCAGAGACAGGTCAAGGAGAGCTGGGAAAGTGGGAGTGGCAGTCACTAGCTTTTCGGGC 2559
QY 510 SerMetGluIleIleGluValSer 517
Db 2560 AGCATGGAATCATTTAGGTCTCC 2583
RESULT 9
ABK49402
ID ABK49402 standard; cDNA; 3544 BP.
XX
AC ABK49402;
XX
XX
XX 02-JUL-2002 (first entry)
XX cDNA encoding human dual specificity phosphatase 21117 protein.
XX
XX Human; dual specificity phosphatase 21117; erythroid-related disorder;
XX haematopoietic-related disorder; leukaemia; autoimmune disease; anaemia;
XX erythrocytosis; liver-related disorder; cancer; gene; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 589..2586
XX /*tag= a
XX /product= "Human dual specificity phosphatase 21117"
XX /note= "Specifically claimed in claim 2"
```

```
PN US2002034807-A1.
XX
PD 21-MAR-2002.
XX
PF 23-MAR-2001; 2001US-00816494.
XX
PR 24-MAR-2000; 2000US-0191858P.
XX
PA (MEYE/) MEYERS R A.
XX
PI Meyers RA;
XX
XX WPI; 2002-351088/38.
DR P-PSDB; AAU79929.
XX
XX New nucleic acids, designated 38692 and 21117, encoding dual specificity
PT phosphatases for treating cell proliferation and differentiation
PT disorders including hematopoietic and erythroid-related disorders and
PT cancers.
XX
XX Claim 2; Fig 1; 76pp; English.
XX
XX The present invention relates to new nucleic acids designated 38692 and
CC 21117 encoding dual specificity phosphatase family members. The nucleic
CC acid, polypeptide encoded by it, and antibody specific for the
CC polypeptide may be used to diagnose and treat haematopoietic-related
CC disorders such as leukaemias and autoimmune diseases, erythroid-related
CC disorders such as anaemias and erythrocytosis, liver-related disorders,
CC and cancers, particularly of the breast, colon, adipose, prostate and
CC lung. The present nucleic acid sequence encodes the human dual
CC specificity phosphatase 21117 protein of the invention, as described
CC above
XX
SQ Sequence 3544 BP; 1000 A; 840 C; 822 G; 882 T; 0 U; 0 Other;
```

Alignment Scores:

Pred. No.:	0	Length:	3544
Score:	488.00	Matches:	488
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	94.39%	Indels:	0
DB:	6	Gaps:	0

US-09-964-277-21 (1-517) x ABK49402 (1-3544)

QY 30 GluLeuMetGlnGlnAsnGlyIleGlyTyValLeuAsnAlaSerAsnThrCysProLys 49

Db 1120 GAGCTGATGCAGCAGATGGATTGGTTATGTTTAAATGCCAGCAATACCTGTCCAAAG 1179

QY 50 ProAspPheIleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGlu 69

Db 1180 CCTGACTTTATCCCGAGTCTCAATTTCTCGGTGCTCCGTGAATCAGACGCTTTTGTGAG 1239

QY 70 LysIleLeuProTrpLeuAspLysSerValAspPheIleGlyLysAlaLysAlaSerAsn 89

Db 1240 AAAATTTTGGCGTGGTGGCAAAATCAGTAGATTTCATTGAGNAAGCAAAAGCCTCCAT 1299

QY 90 GlyCysValLeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAla 109

Db 1300 GGATGTCTTCTAGTGCACTGTTTAGCTGGGATCTCCCGTCCGCCACCATCGTATCGCC 1359

QY 110 TyrlleMetIysArgMetAspMetSerLeuAspGluAlaTyArgPheValLysGluLys 129

Db 1360 TACATCATGAGAGAGTGGACATGTCCTTTAGATGAAGCTTTACAGATTTGTGAAGAAAAA 1419

QY 130 ArgProThrIleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyTrpGluLysLys 149

Db 1420 AGACCTACTATATCTCCAAACTTCAATTTTCTGGCCAACTCTCGGACTATGAGAGAAAG 1479

QY 150 IleLysAsnGlnThrGlyAlaSerGlyProLysSerIysLeuLysLeuHisLeuGlu 169

Db 1480 ATTAAGAACCAGACTGGAGCATCAGGGCCCAAGAGCAAACTCAAGCTGTCTGCACCTGGAG 1539

```
Qy 170 LysProAsnGluProValProAlaValSerGluGlyGlyGlnLysSerGluThrProLeu 189
Db 1540 AAGCCAAATGAACCTGTCCCTCTGTCTCAGAGGTGGACAGAAAAGCAGACGCCCTC 1599
Qy 190 SerProCysAlaAspSerAlaThrSerGluAlaAlaGlyGlnArgProValHisPro 209
Db 1600 AGTCCACCTGTGCCACTCTGCTACTCAGAGGCGCAGACAAAGGCCCGTGCATCCC 1659
Qy 210 AlaSerValProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuVal 229
Db 1660 GCCAGCGTGCACGCGTCCAGCGTGCAGCGTGCCTGCTGTAGAGGACAGCCCGCTGGTA 1719
Qy 230 GlnAlaLeuSerCysLeuHisLeuSerAlaAlaAspArgLeuGluAspSerAsnLysLeuLys 249
Db 1720 CAGCGCTCAGTGGGCTGCACCTGTCCGAGACAGCGCTGGAGACAGCAATAAGCTCAAG 1779
Qy 250 ArgSerPheSerLeuAspLysSerValSerAlaSerMetAlaAlaSerLeu 269
Db 1780 CGTTCCTCTCTCGATATCAAAATCAGTTTCATATTCAGCCAGCATGCGACGCTCTTA 1839
Qy 270 HisGlyPheSerSerSerGluAspAlaLeuGluTyrTyrLysProSerThrThrLeuAsp 289
Db 1840 CATGGCTTCTCTCTCATCAGAAATGCTTGGAAATCTACAAACCTTCCACTACTCTGGAT 1899
Qy 290 GlyThrAsnLysLeuCysGlnPheSerProValGlnGluLeuSerGluGlnThrProGlu 309
Db 1900 GGGACCAACAAGCTATGCCAGTTCCTCCCTGTTCAAGAACTATCGAGCAGACTCCCGAA 1959
Qy 310 ThrSerProAspLysGluGluAlaSerLeuProLysLysLeuGlnThrAlaArgProSer 329
Db 1960 ACCAGTCCCTGATAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2019
Qy 330 AspSerGlnSerLysArgLeuHisSerValArgThrSerSerGlyThrAlaGlnArg 349
Db 2020 GACAGCAGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2079
Qy 350 SerLeuLeuSerProLeuHisArgSerGlySerValGluAspAsnTyrHisThrSerPhe 369
Db 2080 TCCCTTTTATCTCCACTGCATCGAAGTGGAGCGGTGGAGGAGCAATTAACACAGCTTC 2139
Qy 370 LeuPheGlyLeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLeuLys 389
Db 2140 CTTTTCGGCTTTCCACGACGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2199
Qy 390 GlyThrPheSerAspLeuAlaProGlnThrSerThrProSerLeuThrSerSerTrp 409
Db 2200 GGCTGGCACTCGGATATCTTGGCCCCCAGACCTCTACCCCTTCCCTGACACAGAGCTGG 2259
Qy 410 TyrPheAlaThrGluSerSerHisPheTyrSerAlaSerAlaIleTyrGlyGlySerAla 429
Db 2260 TATTTTGGCCACAGAGTCCCTCACACTTCTACTCTGCTCAGCCATCTACGGAGGAGTGCC 2319
Qy 430 SerTyrSerAlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerVal 449
Db 2320 AGTTACTCTGCTTACAGTGCAGCCAGCTGCCACTTGGGAGACCAAGTCTATTCTGTG 2379
Qy 450 ArgArgArgGlnLysProSerAspArgAlaAspSerArgArgSerTrpHisGluGluSer 469
Db 2380 CGCAGGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2439
Qy 470 ProPheGluLysGlnPheLysArgArgSerCysGlnMetGluPheGlyGluSerIleMet 489
Db 2440 CCTTTGAAAGAGGAGTTTAAACGAGAACTGCCAAATGGAAATTTGGAGAGCATCATG 2499
Qy 490 SerGluAsnArgSerArgGluGluLeuGlyLysValGlySerGlnSerSerPheSerGly 509
Db 2500 TCAGAGAAAGGTCAGCGGAGAGGAGTGGGGAAAGTGGGCGAGTCAGTCTAGCTTTTCGGG 2559
Qy 510 SerMetGluIleLeuGluValSer 517
Db 2560 AGCATGAAATCATTAGGTCTCC 2583
```

RESULT 10

```
ABK14474
ID ABK14474 standard; cDNA; 3766 BP.
XX
AC ABK14474;
XX
DT 08-MAY-2002 (first entry)
XX
DE Human protein phosphatase 7 (PP7) cDNA sequence.
XX
KW Human; protein phosphatase 7; PP7; immune system disorder; AIDS;
KW acquired immunodeficiency syndrome; thymic hypoplasia; anaemia; asthma;
KW Crohn's disease; neurological disorder; epilepsy; Huntington's disease;
KW dementia; Parkinson's disease; developmental disorder; Down's syndrome;
KW cell proliferative disorder; cancer; adenocarcinoma; leukaemia; lymphoma;
KW melanoma; myeloma sarcoma; gene; ss.
XX
OS Homo sapiens.
XX
FH Key' Location/Qualifiers
CDS 538..2535
FT /*tag= a
FT /product= "protein_phosphatase_7_(PP7)"
XX
PN WO200210363-A2.
XX
PD 07-FEB-2002.
XX
PF 26-JUL-2001; 2001WO-US023716.
XX
PR 28-JUL-2000; 2000US-0221679P.
PR 03-AUG-2000; 2000US-0223272P.
PR 10-AUG-2000; 2000US-0224309P.
PR 18-AUG-2000; 2000US-0226728P.
PR 30-AUG-2000; 2000US-0229254P.
PR 08-SEP-2000; 2000US-0231366P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Tang YT, Elliott VS, Ramkumar J, Yao MG, Burford N, Wang YE;
PI Stewart EA, Gandhi AR, Patterson C, Lee EA, Hafalia AJA, Lu DAM;
PI Tribouley CM, Griffin JA, Baughn MR, Yue H, Warren BA, Nguyen DB;
PI Walla NK, Kearney L;
XX
DR WPI; 2002-188735/24.
DR P-PSDB; AAU75789.
XX
PT New protein phosphatases, useful for diagnosing, treating or preventing
PT immune system disorders (e.g. Crohn's disease), neurological disorders
PT (e.g. Parkinson's disease), or cell proliferative disorders (e.g.
PT cancers).
XX
PS Claim 5; Page 114-115; 117pp; English.
XX
CC The present invention relates to a new polypeptide, a naturally occurring
CC amino acid sequence at least 95 % identical to it, a biologically active
CC fragment of it or an immunogenic fragment of it. The polypeptides,
CC polynucleotides, agonists and antagonists are useful for diagnosing,
CC treating or preventing disorders associated with aberrant expression of
CC protein phosphatases (PP), particularly immune system disorders e.g.
CC acquired immunodeficiency syndrome (AIDS), thymic hypoplasia, anaemia,
CC asthma or Crohn's disease, neurological disorders e.g. epilepsy,
CC Huntington's disease, dementia or Parkinson's disease, developmental
CC disorders e.g. Down's syndrome, or cell proliferative disorders e.g.
CC cancers including adenocarcinoma, leukaemia, lymphoma, melanoma, myeloma
CC or sarcoma. The present nucleic acid sequence encodes human protein
CC phosphatase 7 (PP7) which is one of several human protein phosphatases
CC (AAU75783-AAU75792) of the invention
XX
SQ Sequence 3766 BP; 1011 A; 903 C; 918 G; 934 T; 0 U; 0 Other;
```

Alignment Scores:  
Pred. No.: 0 Length: 3766  
Score: 488.00 Matches: 488



Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	94.39%	Indels:	0
DB:	6	Gaps:	0
US-09-964-277-21 (1-517) x ABK14474 (1-3766)			
QY	30	GluLeuMetGlnGlnAsnGlyLeuGlyTyrValLeuAsnAlaSerAsnThrCysProLys	49
DB	1069	GAGCTGATGCAGCAAGATGGGATGGTTATGTGTTAAATCCAGCAATACCTGTCCAAAG	1128
QY	50	ProAspPheLeuProGluSerHisPheLeuArgValProValAsnAspSerPheCysGlu	69
DB	1129	CCTGACTTTATCCCGAGTCTCAITTCCTCGGTGCTGCTGTGATGACACTTTGTGAG	1188
QY	70	LysIleLeuProTrpLeuAspLysSerValAspPheIleGluLysAlaLysAlaSerAsn	89
DB	1189	AAAATTTTTCGCGTGTTCGACAAATCAGTAGATTTCATTGAGAAGCAAAAGCTCCCAAT	1248
QY	90	GlyCysValLeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAla	109
DB	1249	GGATGTGTTCTAGTGCACTGTTTAGCTGGGATCTCCCGCTCCGCCACCATCGCTATCGCC	1308
QY	110	TyrIleMetLysArgMetAspMetSerLeuAspGluAlaTyrArgPheValLysGluLys	129
DB	1309	TACATCATGAAGAGATGCACATGCTTTAGATGAAGCTTACAGATTGTGAAAAGAAAA	1368
QY	130	ArgProThrIleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGluLysLys	149
DB	1369	AGACTTACTATCTCCAAACTTCAATTTTCTGGCCCACTCTCTGGACTATGAGAAGAAG	1428
QY	150	IleLysAsnGlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuLeuHisLeuGlu	169
DB	1429	ATTAAGAACCAAGATGGAGCATCAGGCGCAAGAGCAAACTCAAGCTGCTGCACCTGGAG	1488
QY	170	LysProAsnGluProValProAlaValSerGluClyGlyGlnLysSerGluThrProLeu	189
DB	1489	AAGCCAAATGAACCTGCTCCCTGCTGTCTCAGAGGGTGCACAGAAAACGAGACGCCCTC	1548
QY	190	SerProCysAlaAspSerAlaThrSerGluAlaAlaGlyGlnArgProValHisPro	209
DB	1549	AGTCACCTCTGCCGACTCTGCTACTCAGAGGCAGCAGCAAGAGCCCGTGCATCCC	1608
QY	210	AlaSerValProSerValGlnProSerLeuLeuGluAspSerProLeuVal	229
DB	1609	GCCAGCGTCCCGAGCGTCCCGAGCGTGCAGCGTCTGTAGAGGACAGCCCGCTGTA	1668
QY	230	GlnAlaLeuSerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnLysLeuLys	249
DB	1659	CAGGCCCTCAGTGGCTGCACCTGTCGACAGAGGCTGGAAGACAGCAATAAGCTCAAG	1728
QY	250	ArgSerPheSerLeuAspIleLysSerValSerTyrSerAlaSerMetAlaLysLeu	269
DB	1729	CGTTCCCTTCTCTGGATATCAATCAGTTTCAATTCAGCCAGCATGGCAGCATCTTA	1788
QY	270	HisGlyPheSerSerGluAspAlaLeuGluTyrTyrLysProSerThrThrLeuAsp	289
DB	1789	CATGGCTTCTCTCATCAGAGATGCTTTGGAATCTACAAACCTTCCACTCTGGAT	1848
QY	290	GlyThrAsnLysLeuCysGlnPheSerProValGlnGluLeuSerGluGlnThrProGlu	309
DB	1849	GGGACCAACAGCTATGCCAGTTCTCCCTGTTCCAGGAATCTCCGAGCAGATCCCGAA	1908
QY	310	ThrSerProAspLysGluGluAlaSerIleProLysLysLeuGlnThrAlaArgProSer	329
DB	1909	ACCAGTCTGTAAAGGAGAGAGCAGCATCCCAAGAGCTGCAGACCGCCAGCGCTTCA	1968
QY	330	AspSerGlnSerLysArgLeuHisSerValArgThrSerSerGlyThrAlaGlnArg	349
DB	1969	GACAGCCAGACAGCGATTCTGCTCGTACAGACAGAGAGTGGCAGCCGCCAGAGG	2028
QY	350	SerLeuLeuSerProLeuHisArgSerGlySerValGluAspAsnTyrHisThrSerPhe	369
DB			

Db	2029	TCCCTTTTATCTCCACTGCATCGAAGTGGGAGCGTGGAGACAATTACCACACAGCTTC	2088
QY	370	LeuPheGlyLeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLeuLys	389
DB	2089	CTTTTCGGCTTTCCACAGCAGCAGCAGCAGCTCAGAAAGTCTGCTGGCCTGGCCCTAAG	2148
QY	390	GlyTrpHisSerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTrp	409
DB	2149	GGCTGGCAGCTCGGATATCTTTGGGCCCCCAGACCTTACCCCTTCCCTGACCAAGCTGG	2208
QY	410	TyrPheAlaThrGluSerSerHisPheTyrSerAlaSerAlaIleTyrGlyGlySerAla	429
DB	2209	TATTTTCCACAGAGTCTCCTCAGCTTCTACTCTGCTCAGCCATCTACGGAGCAGTGCC	2268
QY	430	SerTyrSerAlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerVal	449
DB	2269	AGTTACTCTGCTACAGCTGCAGCCAGCTGCCACTTGGGAGACCAAGTCTATTCTGTG	2328
QY	450	ArgArgArgGlnLysProSerAspArgAlaAspSerArgSerTrpHisGluGluSer	469
DB	2329	CGCAGCGCGCAGAGCAAGTACAGAGCTGACTCGCGCGGAGCTGGCATGAAGAGAGC	2388
QY	470	ProPheGluLysGlnPheLysArgArgSerCysGlnMetGluPheGlyGluSerIleMet	489
DB	2389	CCCTTTGAAAAGCAGTTTAAACCGCAGAGCTGCCAAATGGAATTTGGAGAGAGCATATG	2448
QY	490	SerGluAsnArgSerArgGluGluLeuGlyLysValGlySerGlnSerSerPheSerGly	509
DB	2449	TCAGAGAACAGGTCACGGAGAGCTGGGGAAGTGGGCAAGTGCAGTCTAGCTTTTCGGGC	2508
QY	510	SerMetGluIleIleGluValSer	517
DB	2509	AGCATGGAAATCATTTAGGTCTCC	2532
RESULT 11			
ABN83966 standard; DNA; 4790 BP.			
ID	ABN83966		
XX	AC	ABN83966;	
XX	AC	ABN83966;	
DT	06-SEP-2002	(first entry)	
XX	XX	Human gene sequence #13.	
DE	Human	gene sequence #13.	
XX	Human	brain; tonsil; hippocampus; foetal brain; diagnosis; gene; ds.	
KW	Human	brain; tonsil; hippocampus; foetal brain; diagnosis; gene; ds.	
XX	Homo sapiens.		
OS	Homo sapiens.		
XX	Key	Location/Qualifiers	
PH	CDS	184..2181	
FT		/*tag= a	
XX	WO200252005-A1.		
PN	04-JUL-2002.		
XX	20-DEC-2001; 2001WO-JP011217.		
PF	22-DEC-2000; 2000JP-00389742.		
XX	(KAZU-) KAZUSA DNA RES INST FOUND.		
PA	(CELE-) CELESTAR LEXICO-SCI LTD.		
XX	Ohara O, Nagase T, Nakajima D;		
PI	WPI; 2002-500762/53.		
XX	P-PSDB; ABB97946.		
DR	Genes and their expression products cloned from human cDNA libraries for		
XX	treatment and diagnosis of diseases associated with their expression.		
PT	Claim 1(a); Page 111-117; 238pp; Japanese.		
XX			

CC The invention relates to DNA encoding polypeptides directly cloned from  
 CC cDNA libraries originating in adult whole brain, human tonsil, human  
 CC adult hippocampus and human foetal whole brain. Polypeptides and  
 CC polynucleotides of the invention may be used in the investigations of  
 CC differential expression of the DNA sequences in normal subjects and  
 CC disease patients. They may also be used in the production of antibodies,  
 CC oligonucleotide probes and DNA chips for diagnosis and identification of  
 CC drugs for treatment of diseases with which the DNA sequences are  
 CC associated. The sequences given in records ABN83954-ABN83984 represent  
 CC human gene sequences of the invention  
 XX

SQ Sequence 4790 BP; 1293 A; 1064 C; 1152 G; 1281 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 0 Length: 4790  
 Score: 488.00 Matches: 488  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 94.39% Indels: 0  
 DB: 6 Gaps: 0

US-09-964-277-21 (1-517) x ABN83966 (1-4790)

Qy 30 GluLeuMetGlnGlnAenglyleGlyTyrValLeuAenAlaSerAsnThrCysProlys 49  
 Db 715 GAGCTGATGAGCAGAAATGGATTGTTATGTTAAATGCCAGCAATACCTGTCCAAAG 774  
 Qy 50 ProAspPheLeuProGluSerHisPheLeuArgValProValAsnAspSerPheCysGlu 69  
 Db 775 CCTGACTTTATCCCGAGTCTCATTTCCCGTGTCCTGTAAGTACAGCTTTGTGAG 834  
 Qy 70 LysileLeuProThrLeuAspLysSerValAspPheLeuGlyLeuAlaLysAlaSerAsn 89  
 Db 835 AAAATTTTGGCGTGGTGGCAATACAGTAGATTTTCATTGAGAAAGCAAAAGCTCCCAAT 894  
 Qy 90 GlyCysValLeuValHisCysleuAlaGlyIleSerArgSerAlaThrIleAlaIleAla 109  
 Db 895 GGAATGTTCTAGTGCACTGTTTAGTGGGATCTCCCGCTCCGCCACCATGCTATCGCC 954  
 Qy 110 TyrIleMetLysArgMetAspMetSerLeuAspGluAlaTyrArgPheValLysGluLys 129  
 Db 955 TACATCATGAGAGGATGGACATGCTTTAGATGAAGCTTACAGATTTGTGAAGAAAA 1014  
 Qy 130 ArgProThrIleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGluLysLys 149  
 Db 1015 AGACCTACTATATCTCAAACTTTCAATTTCTGGGCCAACTCTGGACTATGAGAAGAAG 1074  
 Qy 150 IleLysAenGlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuLeuHisLeuGlu 169  
 Db 1075 ATTAAGAACCCAGACTGGAGCATCAGGGCCAAAGAGCAAACTCAAGCTGCTGACTGGAG 1134  
 Qy 170 LysProAsnGluProValProAlaValSerGluGlyGlnLysSerGluThrProLeu 189  
 Db 1135 AAGCAAAATGAACCTGTCCTGCTGTCTCAGAGGGTGGACAGAAAAAGCGAGAGCCCTC 1194  
 Qy 190 SerProCysAlaAspSerAlaThrSerGluAlaAlaGlyGlnArgProValHisPro 209  
 Db 1195 AGTCCACCTCTGCGGACTCTGACTCTAGAGCAGAGGAAAGGCCCGCTGATCC 1254  
 Qy 210 AlaSerValProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuVal 229  
 Db 1255 GCCAGCTGCCAGCGTCCAGCGTCCAGCGTCTGTTAGAGGACAGCCCGCTGGTA 1314  
 Qy 230 GlnAlaLeuSerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnLysLysLys 249  
 Db 1315 CAGCGCTCAGTGGCGTGCACCTGTCGGCAGACAGGCTGGAAGACAGCAATAGCTCAAG 1374  
 Qy 250 ArgSerPheSerLeuAspIleLysSerValSerTyrSerAlaSerMetAlaAlaSerLeu 269  
 Db 1375 CGTTCCTCTCTGGATATCAATACAGTTTCATATTCAGCCAGCATGCGAGCATCCCTA 1434  
 Qy 270 HisGlyPheSerSerGluAspAlaLeuGluTyrTyrLysProSerThrThrLeuAsp 289

Db 1435 CATGCTTCTCTCATCAGAAAGATGCTTTGGAATACTACAAACCTTCCACTACTCTGGAT 1494  
 Qy 290 GlyThrAenLysLeuCysGlnPheSerProValGlnGlnLeuSerGluGlnThrProGlu 309  
 Db 1495 GGGACCAACAAGCTATGCGAGTCTTCCCTGTTTCCAGGAAGTATCGGAGCAGATCCCGAA 1554  
 Qy 310 ThrSerProAspLysGluGluAlaSerIleProLysLysLeuGlnThrAlaArgProSer 329  
 Db 1555 ACATGCTCTGATAGGAGAGAGCAGCATCCCCAAGAAGCTGCGAGCGCCAGCCTTCA 1614  
 Qy 330 AspSerGlnSerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArg 349  
 Db 1615 GACGCCAGCAGCAGCGATTGCTTCCGTGAGAACCCAGCAGCAGTGGCAGCCGCCAGG 1674  
 Qy 350 SerLeuLeuSerProLeuHisArgSerGlySerValGluAspAsnTyrHisThrSerPhe 369  
 Db 1675 TCCCTTTTATCTCCACTCATCGAAGGGAGCGCTGGAGACAATTACCCAGCAGCTTC 1734  
 Qy 370 LeuPheGlyLeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLeuLys 389  
 Db 1735 CTTTTCGGCTTTCCACCAGCAGCAGCACCTCAGAAAGTCTGCTGGCCTGGGCTTTAAG 1794  
 Qy 390 GlyTyrHisSerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTip 409  
 Db 1795 GGTGGCTGCTCGATATCTTGGCCCCCAGACCTTACCCCTTCCCTGACCCAGCAGCTGG 1854  
 Qy 410 TyrPheAlaThrGluSerSerHisPheTyrSerAlaSerAlaIleTyrGlyGlySerAla 429  
 Db 1855 TATTTTGGCAGAGTCTCTACACTTCTACTCTGCTCAGCCATCTACGGAGGAGTGGC 1914  
 Qy 430 SerTyrSerAlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerVal 449  
 Db 1915 AGTACTCTGCTACAGCTGCGAGCAGCTGCCACTTGGCGAGACCAAGTCTATCTGTG 1974  
 Qy 450 ArgArgGlnLysProSerAspArgAlaAspSerArgArgSerTyrHisGluGluSer 469  
 Db 1975 CGCAGCGCGCAGAACCAAGTACAGAGCTGACTCGCGCGAGCTGGCATGAAGAGAGC 2034  
 Qy 470 ProPheGluLysGlnPheLysArgArgSerCysGlnMetGluPheGlyGluSerIleMet 489  
 Db 2035 CCTTTGAAAGCAGTAAACGCCAAGAGCTGCCAAATGGAAATTTGGAGAGAGCATCATG 2094  
 Qy 490 SerGluAsnArgSerArgGluGluLeuGlyLysValGlySerGlnSerSerPheSerGly 509  
 Db 2095 TAGAGAACAGCTCAGGGAAGAGCTGGGGAAGTGGGCAAGTGGGCAAGTCTAGCTTTTCGGGC 2154  
 Qy 510 SerMetGluIleIleGluValSer 517  
 Db 2155 AGCATGGAATCATTGAGTCTCC 2178

RESULT 12

ABV20833  
 ID ABV20833 standard; cDNA; 5145 BP.  
 XX AC ABV20833;  
 XX 13-SEP-2002 (first entry)  
 DT  
 XX Human prostate expression marker cDNA 20824.

Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
 pharmacogenomic marker; gene; ss.

OS Homo sapiens.

XX W0200160860-A2.

XX 23-AUG-2001.

XX 20-FEB-2001; 2001WO-US005171.

XX 17-FEB-2000; 2000US-0183319P.

XX 16-MAR-2000; 2000US-0189862F.



AC ABV21080;  
XX  
XX  
DT 13-SEP-2002 (first entry)  
XX  
DE Human prostate expression marker cDNA 21071.  
XX  
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
KW pharmacogenomic marker; gene; ss.  
XX  
OS Homo sapiens.  
XX  
XX W0200160860-A2.  
XX  
XX 23-AUG-2001.  
XX  
XX 20-FEB-2001; 2001WO-US005171.  
XX  
XX 17-FEB-2000; 2000US-0183319P.  
XX  
XX 16-MAR-2000; 2000US-0189862P.  
XX  
XX 25-MAY-2000; 2000US-0207454P.  
XX  
XX 09-JUN-2000; 2000US-0211314P.  
XX  
XX 18-JUL-2000; 2000US-0219007P.  
XX  
XX 13-DEC-2000; 2000US-0255281P.  
XX  
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
XX  
XX Schlegel R, Endege WO, Monahan JE;  
XX  
XX WPI; 2001-662795/76.  
XX  
XX Novel isolated nucleic acid molecule associated with cancerous state of  
PT prostate cells and correlating with presence of prostate cancer, useful  
PT for detecting presence of prostate cancer, stage of prostate cancer.  
XX  
XX Claim 1; Page 3481; 11750pp; English.  
XX  
XX The invention relates to an isolated nucleic acid molecule (I) comprising  
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
CC specification or its complement. (I) is useful for: (a) assessing whether  
CC a patient is afflicted with prostate cancer; (b) monitoring the  
CC progression of prostate cancer in a patient; (c) assessing the efficacy  
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing  
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)  
CC determining whether prostate cancer has metastasized in a patient; (h)  
CC assessing the aggressiveness or indolence of prostate cancer in a patient  
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker  
XX  
SQ Sequence 5145 BP; 1408 A; 1135 C; 1253 G; 1346 T; 0 U; 3 Other;

Alignment Scores:  
Pred. No.: 0 Length: 5145  
Score: 488.00 Matches: 488  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 94.39% Indels: 0  
DB: 5 Gaps: 0

US-09-964-277-21 (1-517) x ABV21080 (1-5145)

QY 30 GluLeuMetGlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerAsnThrCysProlys 49  
Db 1120 GAGCTGATGCAGCAGAATGGGATTGGTTATGTTAAATGCCAGCAATACCTGTGCCAAAG 1179  
QY 50 ProAspPheIleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGlu 69  
Db 1180 CCTGACTTTATCCCGAGTCTCAATTCCTCGCGTGGCCCTGTGAATGACAGCTTTTGTGAG 1239  
QY 70 LysIleLeuProTirLeuAspLysSerValAspPheIleGluLeuAlaLysAlaSerAsn 89  
Db 1240 AANAATTGCGGTGGTGGACAATCAGTAGATTTCATTGAGAAAGCAAGCCCTCCAAAT 1299

QY 90 GlyCysValLeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAla 109  
Db 1300 GGAATGTTCTAGTGCATCTGTTAGCTGGGATCTCCCGCTCCGCCACCATCGCTATCGCC 1359  
QY 110 TyrIleMetLysArgMetAspMetSerLeuAspGluAlaTyrArgPheValLysGluLys 129  
Db 1360 TACATCATGAAGAGATGACATGTTCTTTAGATGAAGCTTACAGATTTGTGAAGAAAAA 1419  
QY 130 ArgProThrIleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGluLysLys 149  
Db 1420 AGACCTACTATATCTCCAACTTCAATTTTCTGGGCCAACTCTCTGGACTATGAGAAGAG 1479  
QY 150 IleLysAsnGlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuHisGluGlu 169  
Db 1480 ATTAAGAACCAGACTGGAGCATCGGGCCAAAGAGCAAACTCAAGCTGCTGCACCTCGAG 1539  
QY 170 LysProAsnGluProValProAlaValSerGluGlyGlnLysSerGluThrProLeu 189  
Db 1540 AAGCCAAATGAACCTGTCTCTGCTGTCTCAGAGGGTGGACAGAAAAGCGAGACGCCCTC 1599  
QY 190 SerProProCysAlaAspSerAlaThrSerGluAlaAlaGlyGlnArgProValHisPro 209  
Db 1600 AGTCCACCTGTGCGGACTCTGTACTCTCAGAGCGAGCAGGACAAAGGCCGCTGCATCCC 1659  
QY 210 AlaSerValProSerValProSerValGlnProSerIleLeuGluAspSerProLeuVal 229  
Db 1660 GCCAGCGTCCCGAGCGTCCCGAGCGTGCAGCGCTGCTGTATTAGAGGACAGCCCTCGTA 1719  
QY 230 GlnAlaLeuSerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnLysLeuLys 249  
Db 1720 CAGGCGCTCAGTGGGCTGCACCTGTCCGACAGAGCTGGGAAGACAGCAATAAGCTCAAG 1779  
QY 250 ArgSerPheSerLeuAspIleLysSerValSerTyrSerAlaSerMetAlaAlaSerLeu 269  
Db 1780 CGTTCCTTCTCTCGGATATCAAAATCAGTTTCATATTACAGCCAGCATGGCAGCATCCTTA 1839  
QY 270 HisGlyPheSerSerSerGluAspAlaLeuGluTyrTyrLysProSerThrThrLeuAsp 289  
Db 1840 CATGGCTTCTCTCATCAGAGATGCTTTGGAATACTACAAACCTTCCACTACTCTGGAT 1899  
QY 290 GlyThrAsnLysLeuCysGlnPheSerProValGlnGluLeuSerGluGlnThrProGlu 309  
Db 1900 GGGACCAACAAGCTATGCGAGTTCTCCCTCTTCAGGAACCTATCGGAGCAGACTCCCGAA 1959  
QY 310 ThrSerProAspLysGluGluAlaSerIleProLysLysLeuGlnThrAlaArgProSer 329  
Db 1960 ACCAGTCTGTATGAAGAGAGAGCCAGCATCCCAAGAGCTGCAGACCCCGAGCCCTTCA 2019  
QY 330 AspSerGlnSerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArg 349  
Db 2020 GACGCCAGAGCAAGGATTGCAITTCGTCAGAACCCAGCAGCAGTGGCACCAGGAGG 2079  
QY 350 SerLeuLeuSerProLeuHisArgSerGlySerValGluAspAsnTyrHisThrSerPhe 369  
Db 2080 TCCCTTTTATCTCCACTGCATCGAAGTGGGAGCTGGAGGCAATATACCACACAGCTTC 2139  
QY 370 LeuPheGlyLeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLys 389  
Db 2140 CTTTTCGGCTTTCCACCCAGCAGCAGCACCTCAGAAAGTCTGTGGCTGGGCTTAAAG 2199  
QY 390 GlyTrpHisSerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTirp 409  
Db 2200 GGTGGCACTCGGATATCTTGGCCCCCAGACCTCTACCCCTTCCCTGACCAGCAGTGG 2259  
QY 410 TyrPheAlaThrGluSerSerHisPheTyrSerAlaSerAlaIleTyrGlySerAla 429  
Db 2260 TATTTTGCCACAGAGTCTCTACATTTCTACTCTGCTTCAGCCATCTACGAGGAGGAGTGC 2319  
QY 430 SerTyrSerAlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerVal 449  
Db 2320 AGTTACTCTGCCTACAGCTGCAGCGCTGCCACTTGGCGAGACCAAGTCTATTCTGTG 2379  
QY 450 ArgArgGlnLysProSerAspArgAlaAspSerArgSerTirpHisGluGluSer 469



Db	2080	TCCTCTTTATCTCCACATGCGATCGAGTGGGAGCGTGGAGGACAATATCCACACAGCTTC	2139
Qy	370	LeuPheGlyLeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLeuLys	389
Db	2140	CTTTTCGGCCCTTCCACAGCAGCAGCACCTCAGGAAGTCTGCTGGCTGGGCTTTAAG	2199
Qy	390	GlyTyrHisSerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTyr	409
Db	2200	GGCTGGCACTCGGATATCTTTGGCCCCCAGACCTCTACCCCTTCCTGACGAGCAGCTGG	2259
Qy	410	TyrPheAlaThrGluSerSerHisPheTyrSerAlaSerAlaIleTyrGlyGlySerAla	429
Db	2260	TATTTTGGCCACAGAGTCCTCACACTTCTACTCTGCTCAGCCATCTACGGAGGAGTGCC	2319
Qy	430	SerTyrSerAlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerVal	449
Db	2320	AGTTACTCTGCTCTACAGTCGAGCAGCTGCCACCTTGGGAGACCAAGTCTATTCTGTG	2379
Qy	450	ArgArgArgGlnLysProSerAspArgAlaAspSerArgArgSerTyrHisGluGluSer	469
Db	2380	CCACGGCGCAGAGGCCAAGTGCACAGAGCTGACTCGCGCGGAGCTGGCATGAGAGAGC	2439
Qy	470	ProPheGluLysGlnPheLysArgArgSerCysGlnMetGluPheGlyGluSerIleMet	489
Db	2440	CCCTTTGAAAAGCAGTTTAAACGCAGAAAGCTGCCAAATGCAATTTGAGAGCATCATG	2499
Qy	490	SerGluAsnArgSerArgGluGluLeuGlyLysValGlySerGlnSerSerPheSerGly	509
Db	2500	TACAGAGACAGGTACGGGAAAGAGCTGGGAAAAGTGGCAGTCAGTCTAGCTTTTCGGGC	2559
Qy	510	SerMetGluIleIleGluValSer	517
Db	2560	AGCATGGAAATCATTTAGGTCCTCC	2583
RESULT 15			
ABV20978			
ID ABV20978 standard; cDNA; 5145 BP.			
XX	AC	ABV20978;	
XX	XX		
DT	13-SEP-2002	(first entry)	
XX	Human prostate expression marker	cdNA 20969.	
XX	Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;		
KW	pharmacogenomic marker; gene; ss.		
XX			
OS	Homo sapiens.		
XX			
PN	WO200160860-A2.		
XX			
PD	23-AUG-2001.		
XX			
PF	20-FEB-2001; 2001WO-US005171.		
XX			
PR	17-FEB-2000; 2000US-0183319P.		
PR	16-MAR-2000; 2000US-0189862P.		
PR	25-MAY-2000; 2000US-0207454P.		
PR	09-JUN-2000; 2000US-0211314P.		
PR	18-JUL-2000; 2000US-0219007P.		
PR	13-DEC-2000; 2000US-0255281P.		
XX			
PA	(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.		
XX			
PI	Schlegel R, Endege WO, Monahan JE;		
XX			
DR	WPI; 2001-662795/76.		
XX			
PT	Novel isolated nucleic acid molecule associated with cancerous state of		
PT	prostate cells and correlating with presence of prostate cancer, useful		
PT	for detecting presence of prostate cancer, stage of prostate cancer.		
XX			
PS	Claim 1; Page 3451; 11750pp; English.		

XX The invention relates to an isolated nucleic acid molecule (I) comprising  
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
CC specification or its complement. (I) is useful for: (a) assessing whether  
CC a patient is afflicted with prostate cancer; (b) monitoring the  
CC progression of prostate cancer in a patient; (c) assessing the efficacy  
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing  
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)  
CC determining whether prostate cancer has metastasized in a patient; (h)  
CC assessing the aggressiveness or indolence of prostate cancer in a patient  
CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker  
XX  
SQ Sequence 5145 BP; 1408 A; 1135 C; 1253 G; 1346 T; 0 U; 3 Other;  
  
Alignment Scores:  
Pred. No.: 0 Length: 5145  
Score: 488.00 Matches: 488  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 94.39% Indels: 0  
DB: 5 Gaps: 0  
  
US-09-964-277-21 (1-517) x ABV20978 (1-5145)  
  
QY 30 GluLeuMetGlnGlnAnGlyIleGlyTyrValLeuAsnAlaSerAsnThrCysProLys 49  
Db 1120 GAGCTGATGCAGCAGAAATGGATGGTGTATGTGTTAATGCCAGCAATACCTGTCCCAAG 1179  
  
QY 50 ProAspPheIleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGlu 69  
Db 1180 CCGTACTTATCCCGGAGTCTCATTTCTTCGCTGTGCTGTGAATGACAGCTTTTGTGAG 1239  
  
QY 70 LysIleLeuProTyrLeuAspLysSerValAspPheIleGluValAlaAlaSerAsn 89  
Db 1240 AAAATTTCGCGTGGTGGCAAAATCAGTAGATTTCTTTAGAAAGCAAAAGCCTCCCAT 1299  
  
QY 90 GlyCysValLeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAla 109  
Db 1300 GGAATGTGTTCTAGTGCACTGTTTAGTGGATCTCCCGCTCCGCCACCATCGCTATCC 1359  
  
QY 110 TyrIleMetLysArgMetAspMetSerLeuAspGluAlaTyrArgPheValIleGluLys 129  
Db 1360 TACATCATGAAGAGATGGACATGTCTTTAGATGAAGCTTACAGATTTGTGAAGAAAAA 1419  
  
QY 130 ArgProThrIleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGluLysLys 149  
Db 1420 AGACCTACTATATCTCCAACTTCATTTCTTGGGCCAACTCCTGGACTATGAGAAGAAG 1479  
  
QY 150 IleLysAsnGlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuLeuHisLeuGlu 169  
Db 1480 ATTAAGAACCCAGACTGGAGCATCAGGGCCAAAGAGCAAACTCAAGCTGCTGCACCTGGAG 1539  
  
QY 170 LysProAsnGluProValProAlaValSerGluGlyGlyGlnLysSerGluThrProLeu 189  
Db 1540 AAGCCAAATGAACCTGTCCCTGTCTCAGAGGGTGGACAGAAAGCGAGACGCCCTC 1599  
  
QY 190 SerProProCysAlaAspSerAlaThrSerGluAlaAlaGlyIleArgProValHisPro 209  
Db 1600 AGTCCACCTGTGCCGACTCTGCTACCTCAGAGGCAGCAGGACAAAGGCCCGCTGCATCCC 1659  
  
QY 210 AlaSerValProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuVal 229  
Db 1660 GCCAGCTGCCACGCTGCCAGCGTACGCCGTCTGCTGTTAGAGACAGCCCGCTGGTA 1719  
  
QY 230 GlnAlaLeuSerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnLysLeuLys 249  
Db 1720 CAGCGCCTCAGTGGCGCTGCACCTGCTCCGCACAGCAGCTGGAGACAGCAATAAGCTCAAG 1779  
  
QY 250 ArgSerPheSerLeuAspIleLysSerValSerTyrSerAlaSerMetAlaAlaSerLeu 269  
Db 1780 CGTCTCTCTCTCGGATATCAATCAGTTTCATATTATCCAGCAGCATGCAGCATCTTTA 1839

QY 270 HisGlyPheSerSerSerGluAspAlaLeuGluTyrTyrLysProSerThrThrLeuAsp 289  
Db 1840 CATGGCTTCTCTCATCAGAGATCTTTGGAAATACTACAAACCTTCCACTACTCTGGAT 1899  
QY 290 GlyThrAsnLysLeuCysGlnPheSerProValGlnGluLeuSerGluGlnThrProGlu 309  
Db 1900 GGGACCAACAAGCTATGCCAGTTCTCCCTGTTTCAGGAACCTATCGGAGCAGACTCCCGAA 1959  
QY 310 ThrSerProAspLysGluGluAlaSerIleProLysLysLeuGlnThrAlaArgProSer 329  
Db 1960 ACCAGTCTCTGATAGAGAGAGAGCCAGCATCCCAAGAGCTGCAGACCGCCAGCCCTTCA 2019  
QY 330 AspSerGlnSerLysArgLeuHisSerValArgThrSerSerSerSerGlyThrAlaGlnArg 349  
Db 2020 GACAGCCAGAGACGGATTGCAATTCGGTTCAGACCCAGCAGTGGACCGCCAGAGG 2079  
QY 350 SerLeuLeuSerProLeuHisArgSerGlySerValGluAspAsnTyrHisThrSerPhe 369  
Db 2080 TCCCTTTTATCTCCACTGCATCGAAGTGGAGCGGTGGAGGACAATTACCAACCCAGCTTC 2139  
QY 370 LeuPheGlyLeuSerThrSerClnGlnHisLeuThrLysSerAlaGlyLeuGlyLeuLys 389  
Db 2140 CTTTTCGGCCITTCACCCAGCCAGCAGCACCCTCACGAAGTCTGCTGGCCCTGGGCCCTTAAG 2199  
QY 390 GlyTyrHisSerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTyr 409  
Db 2200 GGCTGGCACTCGGATATCTTGGCCCCCAGACCTCTACCCCTTCCCTGACCCAGCAGCTGG 2259  
QY 410 TyrPheAlaThrGluSerSerHisPheTyrSerAlaSerAlaIleTyrGlyGlySerAla 429  
Db 2260 TATTTTGGCCACAGAGTCTCTCACCTTCTACTCTCCCTCAGCCATCTACGGAGGAGTGGC 2319  
QY 430 SerTyrSerAlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerVal 449  
Db 2320 AGTTACTCTGCTACAGCTGACGCCAGCTGCCCACTTGGGAGACCAAGTCTATTCTGTG 2379  
QY 450 ArgArgArgGlnLysProSerAspArgAlaAspSerArgArgSerTrpHisGluGluSer 469  
Db 2380 CGCAGCGGCGAGAAAGCCCAAGTGCAGAGAGTGACTCGCGCGGAGCTGGCATGAAGAGAGC 2439  
QY 470 ProPheGluLysGlnPheLysArgArgSerCysGlnMetGluPheGlyGluSerIleMet 489  
Db 2440 CCCTTTGAAGAAGCAGTTTAAACGCGAGAGCTGCCAAATGGAAATTTGGAGAGCATCATG 2499  
QY 490 SerGluAsnArgSerArgGluGluLeuGlyLysValGlySerGlnSerSerPheSerGly 509  
Db 2500 TCAGAGAACAGGTCAGCGGAAGAGCTGGGGAAGTGGGCAGTCAGTCTAGCTTTTCGGGC 2559  
QY 510 SerMetGluIleIleGluValSer 517  
Db 2560 AGCATGGAAATCATGTAGGCTCTCC 2583

Search completed: February 28, 2004, 01:26:16  
Job time : 621 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 27, 2004, 18:43:58 ; Search time 222 seconds

(without alignments)

8329.263 Million cell updates/sec

Title: US-09-964-277-20

Perfect score: 3332

Sequence: 1 gagagaaggagaataata.....ataaagatgaactgtttc 3332

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.\*

1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq.\*

2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq.\*

3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq.\*

4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq.\*

5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq.\*

6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2950	88.5	3544	4	US-09-816-494-1
2	1660	49.8	1998	4	US-09-816-494-3
3	247	7.4	279	4	US-09-016-434-91
4	225.2	6.8	2377	4	US-09-920-668-3
5	109.6	3.3	1830	4	US-09-557-921-1
6	95	2.9	2303	4	US-09-922-146-3
7	94.6	2.8	1208	4	US-09-023-655-347
8	90.2	2.7	2109	4	US-09-016-434-1135
9	90.2	2.7	2109	4	US-09-023-655-946
10	86.2	2.6	240	4	US-09-016-434-776
11	85.6	2.6	1619	4	US-09-702-705-801
12	85.6	2.6	1619	4	US-09-736-457-801
13	85.6	2.6	1619	4	US-09-614-124B-801
14	85.6	2.6	1619	4	US-09-671-325-801
15	85.6	2.6	1619	4	US-09-589-184-801
16	85.6	2.6	4637	4	US-09-702-705-804
17	85.6	2.6	4637	4	US-09-736-457-804
18	85.6	2.6	4637	4	US-09-614-124B-804
19	85.6	2.6	4637	4	US-09-671-325-804
20	85.6	2.6	4637	4	US-09-589-184-804
21	84	2.5	1238	2	US-08-530-290-11
22	84	2.5	1238	4	US-09-702-705-803
23	84	2.5	1238	4	US-09-736-457-803
24	84	2.5	1238	4	US-09-614-124B-803
25	84	2.5	1238	4	US-09-671-325-803
26	84	2.5	1238	4	US-09-589-184-803
27	84	2.5	2064	4	US-09-702-705-825

28	84	2.5	2064	4	US-09-736-457-825	Sequence 825, App
29	84	2.5	2064	4	US-09-614-124B-825	Sequence 825, App
30	84	2.5	2064	4	US-09-871-325-825	Sequence 825, App
31	84	2.5	2064	4	US-09-589-184-825	Sequence 825, App
32	84	2.5	2109	4	US-09-702-705-826	Sequence 826, App
33	84	2.5	2109	4	US-09-736-457-826	Sequence 826, App
34	84	2.5	2109	4	US-09-614-124B-826	Sequence 826, App
35	84	2.5	2109	4	US-09-671-325-826	Sequence 826, App
36	84	2.5	2109	4	US-09-589-184-826	Sequence 826, App
37	84	2.5	2240	4	US-09-016-434-1100	Sequence 1100, App
38	82	2.5	2000	4	US-09-016-434-1291	Sequence 1291, App
39	81.2	2.4	1993	2	US-08-990-379-1	Sequence 1, Appli
40	81.2	2.4	1993	2	US-08-990-379-2	Sequence 2, Appli
41	80.4	2.4	944	4	US-09-371-671B-10	Sequence 10, Appli
42	71.6	2.1	539	4	US-09-389-681-311	Sequence 311, App
43	71.6	2.1	539	4	US-09-620-405B-311	Sequence 311, App
44	71.6	2.1	539	4	US-09-339-338-311	Sequence 311, App
45	71.6	2.1	539	4	US-09-433-826B-311	Sequence 311, App

ALIGNMENTS

RESULT 1  
US-09-816-494-1  
; Sequence 1, Application US/09816494  
; Patent No. 6664089

; GENERAL INFORMATION:

; APPLICANT: Meyers, Rachel A.

; TITLE OF INVENTION: 38692 AND 21117, NOVEL DUAL SPECIFICITY

; FILE REFERENCE: 10448-030002

; CURRENT APPLICATION NUMBER: US/09/816,494

; CURRENT FILING DATE: 2001-03-23

; PRIOR APPLICATION NUMBER: US 60/191,858

; PRIOR FILING DATE: 2000-03-24

; NUMBER OF SEQ ID NOS: 10

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1

; LENGTH: 3544

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (589)...(2583)

US-09-816-494-1

Query Match	88.5%	Score 2950;	DB 4;	Length 3544;
Best Local Similarity	95.0%;	Pred. No. 0;		
Matches 3135;	Conservative 0;	Mismatches 0;	Indels 165;	Gaps 2;
QY	197	CGTTTCAGTCAGTGTAAAGCTGTTGGAGCGCGGACCAAGTAAGTAAGTGAATG	256	
DB	224	CGTTTCAGTCAGTGTAAAGCTGTTGGAGCGCGGACCAAGTAAGTAAGTGAATG	283	
QY	257	CGCTGGCTGCTCCAAAGCATCTTTTGTGGAAATGGTATTCCAGTCATCTCTTTATGA	316	
DB	284	CGCTGGCTGCTCCAAAGCATCTTTTGTGGAAATGGTATTCCAGTCATCTCTTTATGA	343	
QY	317	ATCAAAATGTAGGGGCTGCTTTGTGGAGCGAGTCTTTTGAAGACACATCAACGGGAAA	376	
DB	344	ATCAAAATGTAGGGGCTGCTTTGTGGAGCGAGTCTTTTGAAGACACATCAACGGGAAA	403	
QY	377	GAGAAAGACATTCACCTTGGAGGCTCTTGTGAAATGGGTTTAACTCTCTTTTGGC	436	
DB	404	GAGAAAGACATTCACCTTGGAGGCTCTTGTGAAATGGGTTTAACTCTCTTTTGGC	463	
QY	437	AGTCACCACCAAGCTGACCTCATACACTTTTGTAGCAATGGAGTGGCTGAGCTTTGAGC	496	
DB	464	AGTCACCACCAAGCTGACCTCATACACTTTTGTAGCAATGGAGTGGCTGAGCTTTGAGC	523	
QY	497	ACACCAACCTTACATCATCTGCGCAATTAAGAGAGGTGGGAAAGAGGACTTTATG	556	

Db 524 ACACCACCATACATCGTGCACAAATTAAGAGAGAGGTGGGAAAGAGCACTTATTG 583  
Qy 557 TTGTCATGGCCCATGAGATGATGGAACCAAAATGTTACTGAGAGGTGTTGGCTGCG 616  
Db 584 TTGTCATGGCCCATGAGATGATGGAACCAAAATGTTACTGAGAGGTGTTGGCTGCG 643  
Qy 617 TGAAGAATGGAACGGAAGAAAGTCTGCTAAATGATAGCCGGCCATTGTTGGAATACAATA 676  
Db 644 TGAAGAATGGAACGGAAGAAAGTCTGCTAAATGATAGCCGGCCATTGTTGGAATACAATA 703  
Qy 677 CATCCACACATTTTGAAGGCATTAATATCAATCAATGCTCCAAAGCTTATGAAGCGAAGTTGC 736  
Db 704 CATCCACACATTTTGAAGGCATTAATATCAATCAATGCTCCAAAGCTTATGAAGCGAAGTTGC 763  
Qy 737 AACAGACAAAGTGTAAATACAGAGCTCATCCAGCAATTCAGCGAAACATAAAGTTGACA 796  
Db 764 AACAGACAAAGTGTAAATACAGAGCTCATCCAGCAATTCAGCGAAACATAAAGTTGACA 823  
Qy 797 TTGATTCAGTCAGAAAGTTGTAGTTTACGATCAAAAGCTCCCAAGATGTTCCCTCTCT 856  
Db 824 TTGATTCAGTCAGAAAGTTGTAGTTTACGATCAAAAGCTCCCAAGATGTTCCCTCTCT 883  
Qy 857 CTTTCAGACTGTTTCTCACTGACTTCTGGGTAAACTGGAGAGAGCTTCAACTCTGTC 916  
Db 884 CTTTCAGACTGTTTCTCACTGACTTCTGGGTAAACTGGAGAGAGCTTCAACTCTGTC 943  
Qy 917 ACCTGCTTGC----- 926  
Db 944 ACCTGCTTGCAGGTGGTGTGCTGAGTCTCTCGTGTGTTTCCCTGGCCTCTGTGAAGAA 1003  
Qy 927 ----- 926  
Db 1004 AATCCACTCTAGTCCCTACCTGCATTTCTCAGCCTTGTCTTACCTGTGCCAACATGGGC 1063  
Qy 927 -----AGGAGC 932  
Db 1064 CAACCCGAATCTTCCCAATCTTATCTTGGCTGCCAGCGAGATGCTCTCAACAGAGC 1123  
Qy 933 TGATGAGCAGAAATGGATGTTGTTATGTTTAAATGCGAGCAATACCTGTCCAAAGCCTG 992  
Db 1124 TGATGAGCAGAAATGGATGTTGTTATGTTTAAATGCGAGCAATACCTGTCCAAAGCCTG 1183  
Qy 993 ACTTTATCCCGAGTCTCACTTCTCCGCTGCTGCTGTAATCACAGCTTTGTGAGAAA 1052  
Db 1184 ACTTTATCCCGAGTCTCACTTCTCCGCTGCTGCTGTAATCACAGCTTTGTGAGAAA 1243  
Qy 1053 TTTTGGCGTGGTTGGAACAAATCAGTAGATTTCAATTGAGAAAGCAAAAGCCTCCAAATGGAT 1112  
Db 1244 TTTTGGCGTGGTTGGAACAAATCAGTAGATTTCAATTGAGAAAGCAAAAGCCTCCAAATGGAT 1303  
Qy 1113 GTGTTCTAGTGCATGTTTAGCTGGGATCTCCGCTCCGCGACCATCGCTATCGCCTACA 1172  
Db 1304 GTGTTCTAGTGCATGTTTAGCTGGGATCTCCGCTCCGCGACCATCGCTATCGCCTACA 1363  
Qy 1173 TCATGAAGAGGATGGACATGCTTTTAGATGAAGCTTTACAGATTTGTGAAAGAAAAGAC 1232  
Db 1364 TCATGAAGAGGATGGACATGCTTTTAGATGAAGCTTTACAGATTTGTGAAAGAAAAGAC 1423  
Qy 1233 CTACTATATCTCCAACTTCAATTTTCTGGGCCAACTCCTGGACTATGAGAAAGATTA 1292  
Db 1424 CTACTATATCTCCAACTTCAATTTTCTGGGCCAACTCCTGGACTATGAGAAAGATTA 1483  
Qy 1293 AGAACAGACTGGAGATCAGGCGCCAAAGAGCAAACTCAAGCTGTGCACTTGGAGAGC 1352  
Db 1484 AGAACAGACTGGAGATCAGGCGCCAAAGAGCAAACTCAAGCTGTGCACTTGGAGAGC 1543  
Qy 1353 CAAATGAACCTGTCCTGCTGCTCTCAGAGGTGGACAGAAAAGCGAGAGCGCCCTCAGTC 1412  
Db 1544 CAAATGAACCTGTCCTGCTGCTCTCAGAGGTGGACAGAAAAGCGAGAGCGCCCTCAGTC 1603  
Qy 1413 CACCTGTGCGGACTGCTGCTACCTCAGAGCGCAGGACAAAGGCCGCTGATCCCGCA 1472  
Db 1604 CACCTGTGCGGACTGCTGCTACCTCAGAGCGCAGGACAAAGGCCGCTGATCCCGCA 1663

Qy 1473 GGTGCCCAAGCGTGGCCAGCGTGCAGCGTGCCTGCTGTTAGAGGACAGCCCGCTGCTACAGG 1532  
Db 1664 GGTGCCCAAGCGTGGCCAGCGTGCAGCGTGCCTGCTGTTAGAGGACAGCCCGCTGCTACAGG 1723  
Qy 1533 CGCTCAGTGGGCTGCACCTGCTCCGAGACAGGCTGGAAGACAGCAATAAGCTCAAGCGTT 1592  
Db 1724 CGCTCAGTGGGCTGCACCTGCTCCGAGACAGGCTGGAAGACAGCAATAAGCTCAAGCGTT 1783  
Qy 1593 CTTCTCTCTGATATCAAAATCAGTTTCATATTCAGCCAGCATGGCAGCATCCTTACATG 1652  
Db 1784 CTTCTCTCTGATATCAAAATCAGTTTCATATTCAGCCAGCATGGCAGCATCCTTACATG 1843  
Qy 1653 GTTCTCTCTCATCAGAGATGCTTTGGAACTACTACAAACCTTCCACTACTCTGATGGGA 1712  
Db 1844 GTTCTCTCTCATCAGAGATGCTTTGGAACTACTACAAACCTTCCACTACTCTGATGGGA 1903  
Qy 1713 CCAACAAGCTATGCCAGTTCCTCCCTGTTTCAGGAACCTATTCGGAGCAGACTCCCGAAACCA 1772  
Db 1904 CCAACAAGCTATGCCAGTTCCTCCCTGTTTCAGGAACCTATTCGGAGCAGACTCCCGAAACCA 1963  
Qy 1773 GTCTGTATAAGGAGGAAGCCAGCATCCCAAGAGCTGCAGACCGCAGGCTTTCAGACA 1832  
Db 1964 GTCTGTATAAGGAGGAAGCCAGCATCCCAAGAGCTGCAGACCGCAGGCTTTCAGACA 2023  
Qy 1833 GCCAGAGCAAGCGATTGCAATTCGGTTCAGAACCCAGCAGCAGTGGCACCCGCCAGAGTCCC 1892  
Db 2024 GCCAGAGCAAGCGATTGCAATTCGGTTCAGAACCCAGCAGCAGTGGCACCCGCCAGAGTCCC 2083  
Qy 1893 TTTTATCTCCACTGATCGAAGTGGAGCGTGGAGGCAATTACCAACAGCAGCTTCCTTT 1952  
Db 2084 TTTTATCTCCACTGATCGAAGTGGAGCGTGGAGGCAATTACCAACAGCAGCTTCCTTT 2143  
Qy 1953 TCGGCTTTTCCACAGCCAGCAGCACTTCAGAAAGTCTGCTGGCTCGGCTTAAAGGCT 2012  
Db 2144 TCGGCTTTTCCACAGCCAGCAGCAGCACTTCAGAAAGTCTGCTGGCTTAAAGGCT 2203  
Qy 2013 GGCACCTGGATATCTTGGCCGCCAGACCTCTACCCCTTCCCTGACCAAGCAGCTGTATT 2072  
Db 2204 GGCACCTGGATATCTTGGCCGCCAGACCTCTACCCCTTCCCTGACCAAGCAGCTGTATT 2263  
Qy 2073 TTGCACACAGAGTCTCACACTTCTACTCTGCTCAGCCATCTACGGAGGAGTCCAGT 2132  
Db 2264 TTGCACACAGAGTCTCACACTTCTACTCTGCTCAGCCATCTACGGAGGAGTCCAGT 2323  
Qy 2133 ACTTGCCTTACAGCTGCAGCCAGCTGCCCACTTGGGAGAGCCAAAGTCTATTCTGTGGCA 2192  
Db 2324 ACTTGCCTTACAGCTGCAGCCAGCTGCCCACTTGGGAGAGCCAAAGTCTATTCTGTGGCA 2383  
Qy 2193 GCGCGCAGAAAGCCAAAGTGACAGAGCTGACTCGCGCGGAGCTGGCATGAAGAGCCCT 2252  
Db 2384 GCGCGCAGAAAGCCAAAGTGACAGAGCTGACTCGCGCGGAGCTGGCATGAAGAGCCCT 2443  
Qy 2253 TTGAAAGCAGTTTAAACGCGAAGCTGCCAAATTTGGAATTTGGAGAGAGCATCATGTCAG 2312  
Db 2444 TTGAAAGCAGTTTAAACGCGAAGCTGCCAAATTTGGAATTTGGAGAGAGCATCATGTCAG 2503  
Qy 2313 AGAACAGGTCACGGAGAGCTGGGGAAGTGGGAGTCACTGCTAGCTTTTCGGGCGAGCA 2372  
Db 2504 AGAACAGGTCACGGAGAGCTGGGGAAGTGGGAGTCACTGCTAGCTTTTCGGGCGAGCA 2563  
Qy 2373 TGGAAATCAATTGAGTCTCTCTGAGAAAGAAAGACACTTGTGACTTCTATAGACAAATTTTT 2432  
Db 2564 TGGAAATCAATTGAGTCTCTCTGAGAAAGAAAGACACTTGTGACTTCTATAGACAAATTTTT 2623  
Qy 2433 TTTCTTGTTCACAAAAAATTTCCCTGTAAATCTGAAATATATATATATGATACATATAT 2492  
Db 2624 TTTCTTGTTCACAAAAAATTTCCCTGTAAATCTGAAATATATATATATGATACATATAT 2683  
Qy 2493 ATTTTGGAAAAATGAGCTATGCTGTAAAAAGCAACAGGTGGATCAACCCAGTCTGTACTC 2552  
Db 2684 ATTTTGGAAAAATGAGCTATGCTGTAAAAAGCAACAGGTGGATCAACCCAGTCTGTACTC 2743

QY 2553 TCTTAACATCTGCATTTGAGAGATCAGTAACTCTCTCAACAAATGGAGGGAG 2612  
Db 2744 TCTTAACATCTGCATTTGAGAGATCAGTAACTCTCTCAACAAATGGAGGGAG 2803  
QY 2613 ATGTAGAAATCCCCCTAGACGAGGAGAAACCAATTTATTCAGTGAATACACATCCTCT 2672  
Db 2804 ATGTAGAAATCCCCCTAGACGAGGAGAAACCAATTTATTCAGTGAATACACATCCTCT 2863  
QY 2673 TGTTCCTTAAAGCAAGTGTCTTTGGTGTGGAGGACAAATCCCCCTAGCACTTTCCAC 2732  
Db 2864 TGTTCCTTAAAGCAAGTGTCTTTGGTGTGGAGGACAAATCCCCCTAGCACTTTT-CAC 2922  
QY 2733 GTTGTGCTACTAAGAGATCTCAATATATTAGTCTTTGTCGGAGCCCTTCCATAGTACACCT 2792  
Db 2923 GTTGTGCTACTAAGAGATCTCAATATATTAGTCTTTGTCGGAGCCCTTCCATAGTACACCT 2982  
QY 2793 TAGGCTGAGACTGAGCAGGCTTGGGGTCAAGTGTAGGTAGACCTGTGTAGGAGACAGGCC 2852  
Db 2983 TAGGCTGAGACTGAGCAGGCTTGGGGTCAAGTGTAGGTAGACCTGTGTAGGAGACAGGCC 3042  
QY 2853 TAGTGGTAAATCCAGAGAAATGATCTATCCAAAGCTGATTCACAAACCCAGCTCACC 2912  
Db 3043 TAGTGGTAAATCCAGAGAAATGATCTATCCAAAGCTGATTCACAAACCCAGCTCACC 3102  
QY 2913 TCACAGCCGAGGAGACAGAGCATCACTCTGCTGGACGACCATTAGGGGCTTTGCCAAGG 2972  
Db 3103 TCACAGCCGAGGAGACAGAGCATCACTCTGCTGGACGACCATTAGGGGCTTTGCCAAGG 3162  
QY 2973 TCTACCTTAGAGAAACCCAGTACTCAGACAGGAAAGTCGGGGCTTTGACNACTACCAT 3032  
Db 3163 TCTACCTTAGAGAAACCCAGTACTCAGACAGGAAAGTCGGGGCTTTGACNACTACCAT 3222  
QY 3033 ATCTGGTAGCCCATTTCTAGGCATTTGTAATAGTGTAGGTAGTCTACACTTTTCAGA 3092  
Db 3223 ATCTGGTAGCCCATTTCTAGGCATTTGTAATAGTGTAGGTAGTCTACACTTTTCAGA 3282  
QY 3093 CCAATTCAAATCTGTATGACACAAATTCCTGGGGCTAGATGGAGATAATTTTTTTTT 3152  
Db 3283 CCAATTCAAATCTGTATGACACAAATTCCTGGGGCTAGATGGAGATAATTTTTTTTT 3342  
QY 3153 CTTCTCAGCTTTATGAGAGAGGAACTGTCTAGGATTCAGCTGACCAACCAAGAAC 3212  
Db 3343 CTTCTCAGCTTTATGAGAGAGGAACTGTCTAGGATTCAGCTGACCAACCAAGAAC 3402  
QY 3213 TGGCAACATCAGATTTAAAGTAAAGTGGAGGCTAACGAGTCTACCTCCCTCTTTGTA 3272  
Db 3403 TGGCAACATCAGATTTAAAGTAAAGTGGAGGCTAACGAGTCTACCTCCCTCTTTGTA 3462  
QY 3273 AATCAAGAAATTTTAAATGGGAATGTCAATCTTTAAATAAAGATGAACTTGGTTTC 3332  
Db 3463 AATCAAGAAATTTTAAATGGGAATGTCAATCTTTAAATAAAGATGAACTTGGTTTC 3522

RESULT 2  
US-09-816-494-3  
; Sequence 3, Application US/09816494  
; Patent No. 6664089  
; GENERAL INFORMATION:  
; APPLICANT: Meyers, Rachel A.  
; TITLE OF INVENTION: 38692 And 21117, NOVEL DUAL SPECIFICITY  
; TITLE OF INVENTION: PHOSPHATASE MOLECULES AND USES THEREFOR  
; FILE REFERENCE: 10448-030002  
; CURRENT APPLICATION NUMBER: US/09/816,494  
; CURRENT FILING DATE: 2001-03-23  
; PRIOR FILING DATE: 2000-03-24  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 1998  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-816-494-3

Query Match 49.8%; Score 1660; DB 4; Length 1998;  
Best Local Similarity 91.8%; Pred. No. 0;  
Matches 1834; Conservative 0; Mismatches 0; Indels 164; Gaps 1;  
QY 562 ATGCCCATCAGATGATTGGAACTCAAAATTTGTTACTGAGAGGTTGGTGGTCTCTCTGGAA 621  
Db 1 ATGCCCATCAGATGATTGGAACTCAAAATTTGTTACTGAGAGGTTGGTGGTCTCTCTGGAA 60  
QY 622 AGTGGAAACGGAAGAGTGTCTGTAATTTAGTACGGGCCATTTGTGGAAATACATATCATCC 681  
Db 61 AGTGGAAACGGAAGAGTGTCTGTAATTTAGTACGGGCCATTTGTGGAAATACATATCATCC 120  
QY 682 CACATTTTGAAGCCATTAATATCAACTGCTCCAAAGCTTATGAAGCGAAGGTTGCAACAG 741  
Db 121 CACATTTTGAAGCCATTAATATCAACTGCTCCAAAGCTTATGAAGCGAAGGTTGCAACAG 180  
QY 742 GACAAAGTGTAAATTAACAGAGCTCATCCAGCATTTAGCGAAACATTAAGGTTGACATTTGAT 801  
Db 181 GACAAAGTGTAAATTAACAGAGCTCATCCAGCATTTAGCGAAACATTAAGGTTGACATTTGAT 240  
QY 802 TGCAGTCAGAGGTTGTAGTTTACGATCAAGCTCCAGAGATGTTGCCCTCTCTCTCTTCA 861  
Db 241 TGCAGTCAGAGGTTGTAGTTTACGATCAAGCTCCAGAGATGTTGCCCTCTCTCTCTTCA 300  
QY 862 GACTGTTTTCTCACTGTACTTCTGGGTAAACTGGAGAGAGGTTCAACTCTCTCTCACCTG 921  
Db 301 GACTGTTTTCTCACTGTACTTCTGGGTAAACTGGAGAGAGGTTCAACTCTCTCTCACCTG 360  
QY 922 CTTGTC----- 926  
Db 361 CTTGTCAGGTTGGTGTGCTGATTTCTCTCGTTGTTTCCCTGGCTCTGTGAAGGAAATATCC 420  
QY 927 ----- 926  
Db 421 ACTCTAGTCCTACCTGCAATTTCTCAGCCTTGTCTTACCTGTGTCACCAATTTGGGCCAAC 480  
QY 927 -----AGAGCTGATG 937  
Db 481 CGAATTTCTTCCCAATCTTTATCTTGGCTGCAGCGAGATGTCTCAACAGAGAGTGTATG 540  
QY 938 CAGCAGAAATGGGATTTGGTATGTTTAAATGCCAGCAATACCTGTCCAAAGGCTGACCTTT 997  
Db 541 CAGCAGAAATGGGATTTGGTATGTTTAAATGCCAGCAATACCTGTCCAAAGGCTGACCTTT 600  
QY 998 ATCCCCAGAGTCTCAATTTCTCTGCTGTGCTGTGAATGACAGCTTTTGTGAGAAAATTTTG 1057  
Db 601 ATCCCCAGAGTCTCAATTTCTCTGCTGTGCTGTGAATGACAGCTTTTGTGAGAAAATTTTG 660  
QY 1058 CCGTGTGTGACAAATCAGTAGATTTTCAATTTGAGAAAGCAAAAGCTCCATCGATGTGTT 1117  
Db 661 CCGTGTGTGACAAATCAGTAGATTTTCAATTTGAGAAAGCAAAAGCTCCATCGATGTGTT 720  
QY 1118 CTAGTGCACTGTTTAGCTGGGATCTCCCGCTCCGCCACCATCGCTATCGCTTACATCATG 1177  
Db 721 CTAGTGCACTGTTTAGCTGGGATCTCCCGCTCCGCCACCATCGCTTACATCATG 780  
QY 1178 AAGAGGATGGACATGTCTTTTAGATGAAGCTTACAGATTTGTGAAAGAAAAGAGCCTACT 1237  
Db 781 AAGAGGATGGACATGTCTTTTAGATGAAGCTTACAGATTTGTGAAAGAAAAGAGCCTACT 840  
QY 1238 ATATCTCCAAACTTCAATTTTCTGGGCCAACTCTCGGATGATGAGAGAGATTAAGAC 1297  
Db 841 ATATCTCCAAACTTCAATTTTCTGGGCCAACTCTCGGATGATGAGAGAGATTAAGAC 900  
QY 1298 CAGACTGGAGCATCAGGGCCAAAGACAACTCAAGCTGCTGACCTGGAGAGGCCAAT 1357  
Db 901 CAGACTGGAGCATCAGGGCCAAAGACAACTCAAGCTGCTGACCTGGAGAGGCCAAT 960  
QY 1358 GAACCTGTCTGCTGTCTCAGAGGTTGACAGAAAGCGAGCGCCCTCAGTCCACCC 1417  
Db 961 GAACCTGTCTGCTGTCTCAGAGGTTGACAGAAAGCGAGCGCCCTCAGTCCACCC 1020



GENERAL INFORMATION:  
; APPLICANT: Lex M. Cowsett  
; APPLICANT: Brett P. Monia  
; TITLE OF INVENTION: ANTISENSE MODULATION OF DUAL SPECIFIC PHOSPHATASE 8 EXPRESSION  
; FILE REFERENCE: RTS-0246  
; CURRENT APPLICATION NUMBER: US/09/920,668  
; CURRENT FILING DATE: 2001-08-01  
; NUMBER OF SEQ ID NOS: 49  
; SEQ ID NO 3  
; LENGTH: 2377  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; NAME/KEY: CDS  
; LOCATION: (135)...(2012)  
; US-09-920-668-3

Query Match 6.8%; Score 225.2; DB 4; Length 2377;  
Best Local Similarity 60.6%; Pred. No. 7.7e-60;  
Matches 418; Conservative 0; Mismatches 248; Indels 24; Gaps 2;

2y	927	AGGAGCTGATGACGAGCAGATGGGATTTGGTTATGTTAAATGCCAGCAATACCTGTCCAA	986
2b	670	AGGATCTGATGACGCAAAATGGAATAGCTACCTCAACGCCAGCACTCTGCCCCA	729
2y	987	AGGCTGACTTTATCCCGAGTCTCATTTCTGCTGCTGCTGAATGACAGCTTTTGTG	1046
2b	730	AGGCTGACTTCACTCGGAGAGCGGCTTCATCGGGTCCCCATCAACGCACTACTGTG	789
2y	1047	AGAAATTTTGGCGGTGGGCAAAATCAATAGTATTTATTGAGAAAGCAAAAGCCTCCA	1106
2b	790	AAAACCTGCTGCCCTGGTGGCAAGTCCATCGAGTTTCATCATGATTAAGCCAGCTCTCCA	849
2y	1107	ATGGATGTGTTCTAGTGCACATGTTAGCTGGGATCTCCGCTCCGCCACCATCGTATCG	1166
2b	850	GCTGCCAAGTCATCGTCCATCTGCTGGCTGGCATCTCCGCTCTGCCACCATCGCCATCG	909
2y	1167	CCTACATCATGAGAGGATGGACATCTCTTTAGATGAAGCTTTACAGATTTGTGAAGAAA	1226
2b	910	CCTACATCATGAGACCATGGCATGTCTCCGACGACGCTTACAGGTTCTGTGAAGGACA	969
2y	1227	AAAGACTACTATATCTCCAACTTCATTTCTGGGCCAACTCTGACATATGAGAAGA	1286
2b	970	GGCGCCGCTCCATCTCGCCCACTTCAACTTCTCTGGGCCAGCTGCTGGAGTACGAGCGCA	1029
2y	1287	AGATTAAAGACCACTGAGCAGTCAAGGCAAGAGCAAACTCAAGCTGCTGACCTGG	1346
2b	1030	CGCTGAAGCTGCTGGCGCCCTGCAGGGCGACCCGGCACCCCTCAGGAGC---CCGG	1086
2y	1347	AGAAGCCAAATGAACCTGCTCCCTGTCTGCTAGAGGTTGGAAGAAAGCGAGAGCGCCC	1406
2b	1087	AGCCTCGCCCACTCTGCGCGGGGCGCCGCTGCCACGCGTGCACACCATCTACCTCAG	1146
2y	1407	TCAGTCCACCTGTGCGGACTCTGCTACCTCAGAGCGCAGCAGCAAAAGGCCGCTGATC	1466
2b	1147	AGAGCGCTGCCACAGGAATGCGGCTGCCAGGAGGGCG-----	1185
2y	1467	CCGCCAGCTGCCAGCGTCCAGCGTGCAGCGCTGCTGTAGAGGACAGCCGCGTGG	1526
2b	1186	GCCTGAGCGGGGCGGAGSCCCCGCGCCCGCCACGCCCCCGCGGACGAGCACTGC	1245
2y	1527	TACAGGGCTCAGTGGGCTCAGCTGCTCCGACAGAGCTGGAAGACAGCANTAGTCA	1586
2b	1246	AGCAGGGCTGCGCGGCTGCACTCTCTCGACCGCTGAGGACACTAACCCGCTCA	1305
2y	1587	AGGCTTCTCTCTCTGATATCAAAATCAG	1616
2b	1306	AGCGCTCTCTCTCTGGACATCAAGTCTG	1335

RESULT 5  
US-09-557-921-1  
; Sequence 1, Application US/09557921

Patent No. 6551810  
; GENERAL INFORMATION:  
; APPLICANT: Luche, Ralf M.  
; APPLICANT: Wei, Bo  
; TITLE OF INVENTION: DSP-10 DUAL-SPECIFICITY PHOSPHATASE  
; FILE REFERENCE: 200125.416  
; CURRENT APPLICATION NUMBER: US/09/557,921  
; CURRENT FILING DATE: 2000-04-20  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 1830  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; US-09-557-921-1

Query Match 3.3%; Score 109.6; DB 4; Length 1830;  
Best Local Similarity 61.0%; Pred. No. 1.7e-23;  
Matches 178; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

Qy	1006	GTCTCATTTCTCTGCTGCTGCTGCTGAATGACAGCTTTTGTGAGAAATTTTGGCGTGGTT	1065
Db	1187	GTTCAACTACAAGCGCTGCCAGCCACTGACAGCAACAAAGCAGAACCTGCGGAGTACTT	1246
Qy	1066	GGACAAATCAGTAGATTTTCATTGAGAAAGCAAAAGCCTCCAAATGGATGTCTTAGTSCA	1125
Db	1247	TGAAGAGGCTTTTGTGCTTTCATTGAGAAAGCTCACAGTGTGGAGAGGGCTTCTCATCCA	1306
Qy	1126	CTGTTTAGTGGATCTCCGCTCCGCCACCATCGCTATCGCTACATCATGAAGAGAT	1185
Db	1307	CTGCCAGGCTGGGGTGTCCGCTCCGCCACCATCGCTATCTTGTGTAAGACAC	1366
Qy	1186	GGACATGCTTTTAGATGAAGCTTACAGATTTGTGAAAGAAAAGACCTACTATATCTCC	1245
Db	1367	TCGATGACCATGACTGATGCTTATAATTTGTCAAAGCAACGACCAATTAICTCCCC	1426
Qy	1246	AACTTCAATTTCTGGGCCAACTCTCGGACTATGAGAAAGATTAAGAAC	1297
Db	1427	AAACTTAATTTCTGGGGCAGTTCTAGATTCGAGGAGACCTTAAACAAC	1478

RESULT 6  
US-09-922-146-3  
; Sequence 3, Application US/09922146  
; Patent No. 6566133  
; GENERAL INFORMATION:  
; APPLICANT: Lex M. Cowsett  
; APPLICANT: Brett P. Monia  
; TITLE OF INVENTION: ANTISENSE MODULATION OF DUAL SPECIFIC PHOSPHATASE 9 EXPRESSION  
; FILE REFERENCE: RTS-0252  
; CURRENT APPLICATION NUMBER: US/09/922,146  
; CURRENT FILING DATE: 2001-08-01  
; NUMBER OF SEQ ID NOS: 48  
; SEQ ID NO 3  
; LENGTH: 2303  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; NAME/KEY: CDS  
; LOCATION: (114)...(1268)  
; US-09-922-146-3

Query Match 2.9%; Score 95; DB 4; Length 2303;  
Best Local Similarity 58.1%; Pred. No. 8.1e-19;  
Matches 164; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

Qy	1004	GAGTCTCATTTCTCTGCTGCTGCTGCTGAATGACAGCTTTTGTGAGAAAATTTTGGCGTGG	1063
Db	858	GACTTTCATACAGCAGATCCCATCTCCGACCACTGAGCCAGAACCTGTGCGGTTTC	917
Qy	1064	TTGGACAAATCAGTAGATTTTCATTGAGAAAGCAAAAGCCTCCATGGATGTCTTAGTG	1123
Db	918	TTTCCGAGGCCATTTGAGTTTCAATGAGGCTTTGTCCCAAGAACTGCGGGTGTCTGTC	977



Db	1225	CATTGCTTGGCTGGCAATTAGCCGCTCAGTCACTGTGACTGTGGCTTACCTTATGCGAAG	1284
QY	1184	ATGGACATGCTTTTACATCAAGCTTACAGATTGTGAAAGAAAAAGCACTACTATATCT	1243
Db	1285	CTCAATCTGTCGATGAACGATGCTCTATGACATTTGCAAAATGAATAAATCCACATATCC	1344
QY	1244	CCAAACTTCAATTTTCTGGGCCAACTCCTGGACTATGAGAAGAAGAT	1290
Db	1345	CCCTAACTTCAACTTTCATGGGTGAGCTGTGGACTTCGAGAGGAGCGCT	1391

## US-09-023

RESULT 10  
 US-09-016-434-776  
 ; Sequence 776, Application US/09016434  
 ; Patent No. 6500938  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Janice Au-Young  
 ; APPLICANT: Jeffrey J. Seilhamer  
 ; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING  
 ; TITLE OF INVENTION: PATHWAY GENE EXPRESSION  
 ; NUMBER OF SEQUENCES: 1490  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
 ; STREET: 3174 PORTER DRIVE  
 ; CITY: PALO ALTO  
 ; STATE: CALIFORNIA  
 ; COUNTRY: USA  
 ; ZIP: 94304  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/016,434  
 ; FILING DATE: HERewith  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER:  
 ; FILING DATE:  
 ; CLASSIFICATION:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Zeller, Karen J.  
 ; REGISTRATION NUMBER: 37,071  
 ; REFERENCE/DOCKET NUMBER: PA-0002 US  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (650) 855-0555  
 ; TELEFAX: (650) 845-4166  
 ; INFORMATION FOR SEQ ID NO: 776:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 240 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; IMMEDIATE SOURCE:  
 ; LIBRARY: LIVRBCT01  
 ; CLONE: 480457  
 ; US-09-016-434-776

	Query Match	2.6%;	Score 86.2;	DB 4;	Length 240;
	Best Local Similarity	66.5%;	Prod. No. 8.5e-17;		
	Matches 121;	Conservative	0;	Mismatches 61;	Indels 0;
	Gaps	0;			
QY	1116	TTCTAGTGCACGTGTTTAGCTGGGATCTCCGGCTCGGCCACCATCGCTATCGCTACATCA	1175		
DB	10	TTCTCATCCACTGCCAGGCTGGGGTGTCCCGCTCGNCACCATCGTCATGGTTACTTGA	69		
QY	1176	TGAAGAGGATGACATGTCITTTAGATGAAGCTTACAGATTTGTGAAGAAAAAGACCTA	1235		
DB	70	TGAAGCACANTCGGATGACCATGACTGATGCTTATAAATTTGTCAAAGGCAACGACCAA	129		
QY	1236	CTATATCTCCAAACTTCATTTTCTGGGCCCAACTCTCTGGACTATGAGAAGAAGATTAAAGA	1295		



```
Db 130 TTAATCNCCAAACCTTAACCTTCATGGGCGAGTTGCTAGAGTTGCGAGACACCTAAACA 189
QY 1296 AC 1297
Db 190 AC 191

RESULT 11
US-09-702-705-801
; Sequence 801, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 801
; LENGTH: 1619
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-702-705-801

Query Match 2.6%; Score 85.6; DB 4; Length 1619;
Best Local Similarity 53.6%; Pred. No. 5.7e-16;
Matches 178; Conservative 0; Mismatches 154; Indels 0; Gaps 0;

QY 993 ACTTTATCCCGAGTCTCATTTTCCTGCGTGTCCCTGTGTAATGACAGAGCTTTTGTGAGAAAA 1052
Db 429 ACTTTGAAGGACATATCAGTACAAGTGCATCCCAAGTGAAGATAACACAGGCCGACA 488
QY 1053 TTTTGGCGTGGTGGACAAATCAGTAGATTTTCATTTGAGAAAGCAAAAGCCTCCAATGGAT 1112
Db 489 TCAGTCTCTGTTTCATGGAAGCCATAGATACATCGATGCCGTGAGGACTGCCGTGGGC 548
QY 1113 GTGTTCTAGTGCACTGTTTAGTGGGATCTCCCGCTCCGCCACCATCGCTATCGCTTACA 1172
Db 549 GCGTCTGTTGCACTGCCAGCGGGCATCTCGCGTGGCCACCATCTGCTGGCTTACC 608
QY 1173 TCATGAAGAGATGACATGTCTTTAGATGAGCTTACAGTTGTGAGAAAGAAAAAGAC 1232
Db 609 TGATGATGAAGAAACGGGTGAGGCTGAGGAGGCGCTTCGAGTTCGTTAAGCAGCGCGCA 668
QY 1233 CTACTATCTCCAACTTCATTTCTGGGCACTCTGCACTATGAGAGAGATT 1292
Db 669 GCATTATCTCGCCCACTTCAGCTTCATGGGCGAGCTGCTGAGTTCGAGTCCAGGTGC 728
QY 1293 AGAACAGAGCTGGAGCATCAGGCGCCAAAGAC 1324
Db 729 TGCCACGTCCTGTGCTGGAGGCTGTAGC 760

RESULT 12
US-09-736-457-801
; Sequence 801, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
```

```
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 801
; LENGTH: 1619
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-736-457-801

Query Match 2.6%; Score 85.6; DB 4; Length 1619;
Best Local Similarity 53.6%; Pred. No. 5.7e-16;
Matches 178; Conservative 0; Mismatches 154; Indels 0; Gaps 0;

QY 993 ACTTTATCCCGAGTCTCATTTTCCTGCGTGTCCCTGTGTAATGACAGCTTTTGTGAGAAAA 1052
Db 429 ACTTTGAAGGACATATCAGTACAAGTGCATCCCAAGTGAAGATAACACAGGCCGACA 488
QY 1053 TTTTGGCGTGGTGGACAAATCAGTAGATTTTCATTTGAGAAAGCAAAAGCCTCCAATGGAT 1112
Db 489 TCAGTCTCTGTTTCATGGAAGCCATAGATACATCGATGCCGTGAGGACTGCCGTGGGC 548
QY 1113 GTGTTCTAGTGCACTGTTTAGTGGGATCTCCCGCTCCGCCACCATCGCTATCGCTTACA 1172
Db 549 GCGTCTGTTGCACTGCCAGCGGGCATCTCGCGTGGCCACCATCTGCTGGCTTACC 608
QY 1173 TCATGAAGAGATGACATGTCTTTAGATGAGCTTACAGTTGTGAGAAAGAAAAAGAC 1232
Db 609 TGATGATGAAGAAACGGGTGAGGCTGAGGAGGCGCTTCGAGTTCGTTAAGCAGCGCGCA 668
QY 1233 CTACTATCTCCAACTTCATTTCTGGGCACTTCCTGGGAGCTATGAGAGAGATT 1292
Db 669 GCATTATCTCGCCCACTTCAGCTTCATGGGCGAGCTGCTGAGTTCGAGTCCAGGTGC 728
QY 1293 AGAACAGAGCTGGAGCATCAGGCGCCAAAGAC 1324
Db 729 TGCCACGTCCTGTGCTGGAGGCTGTAGC 760

RESULT 13
US-09-614-124B-801
; Sequence 801, Application US/09614124B
; Patent No. 6630574
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Mannion, Jane
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE REFERENCE: 210121.478C9
; CURRENT APPLICATION NUMBER: US/09/614,124B
; CURRENT FILING DATE: 2001-07-11
; NUMBER OF SEQ ID NOS: 1668
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 801
; LENGTH: 1619
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-614-124B-801
```

Query Match 2.6%; Score 85.6; DB 4; Length 1619;  
Best Local Similarity 53.6%; Pred. No. 5.7e-16;  
Matches 178; Conservative 0; Mismatches 154; Indels 0; Gaps 0;

QY 993 ACTTTATCCCCGAGTCTCATTTCTCGGTGTGCTGTGAATGACAGCTTTTGTGAGAAA 1052  
DB 429 ACTTTGAGGACATCTACGATACAAAGTGCATCCAGTGGAGATACCAACAGCCGCA 488

QY 1053 TTTTCCGCTGTTGGCAAAATCAGTAGATTTTCATGTAGAAAGCAAAAGCTTCAATGGAT 1112  
DB 489 TCAGCTCTGTTTCATGGAAGCCATAGATACATCGATGCGGTGAAGGACTGCCGTGGGC 548

QY 1113 GTGTTCTAGTCACTGTTTACTGGATCTCCCGCTCCGCCACATCGCTATCGCCTACA 1172  
DB 549 GCGTCTGTTGATCTGCGAGGGGCACTTCGCGTGGCCACATCTGCTGGCTTACC 608

QY 1173 TCATGAAGAGATGACATGCTTTTAGATGAAGCTTACAGATTTGTAAGAAAAGAC 1232  
DB 609 TCATGATGAAGAAACGGGTGAGGCTGCGGTGGCCACATCTGCTGGCTTACC 668

QY 1233 CTACTATATCTCCAACTTCAATTTTCTGGCCAACTCTCTGGACTATGAGAAGATTA 1292  
DB 669 GCATTATCTCCCAACTTCAGCTTCATGGGGAGCTGCTGCAGTTTCGAGTCCCAAGTGC 728

QY 1293 AGAACCAAGCTGGAGCATCAGGGCCAAAGAC 1324  
DB 729 TGGCCACGCTCTGTGCTGCGGAGGCTGCTAGC 760

RESULT 14  
US-09-671-325-801  
; Sequence 801, Application US/09671325  
; Patent No. 6667154  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Lodes, Michael A.  
; APPLICANT: Fanger, Gary  
; APPLICANT: Vedvick, Tom  
; APPLICANT: Carter, Darrick  
; APPLICANT: Mannion, Jane  
; APPLICANT: Fan, Liqun  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.478C12  
; CURRENT APPLICATION NUMBER: US/09/671.325  
; CURRENT FILING DATE: 2000-09-26  
; NUMBER OF SEQ ID NOS: 1825  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 801  
; LENGTH: 1619  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-671-325-801

Query Match 2.6%; Score 85.6; DB 4; Length 1619;  
Best Local Similarity 53.6%; Pred. No. 5.7e-16;  
Matches 178; Conservative 0; Mismatches 154; Indels 0; Gaps 0;

QY 993 ACTTTATCCCCGAGTCTCATTTCTCGGTGTGCTGTGAATGACAGCTTTTGTGAGAAA 1052  
DB 429 ACTTTGAGGACATCTACGATACAAAGTGCATCCAGTGGAGATACCAACAGCCGCA 488

QY 1053 TTTTCCGCTGTTGGCAAAATCAGTAGATTTTCATGTAGAAAGCAAAAGCTTCAATGGAT 1112  
DB 489 TCAGCTCTGTTTCATGGAAGCCATAGATACATCGATGCGGTGAAGGACTGCCGTGGGC 548

QY 1113 GTGTTCTAGTCACTGTTTACTGGATCTCCCGCTCCGCCACATCGCTATCGCCTACA 1172  
DB 549 GCGTCTGTTGATCTGCGAGGGGCACTTCGCGTGGCCACATCTGCTGGCTTACC 608

QY 1173 TCATGAAGAGATGACATCTTTTAGATGAAGCTTACAGATTTGTAAGAAAAGAC 1232  
DB 609 TGATGATGAAGAAACGGGTGAGGCTGAGGAGGCTTCGAGTTTCGTTAAGCAGCCGCA 668

QY 1233 CTACTATATCTCCAACTTCAATTTTCTGGCCAACTCTCTGGACTATGAGAAGATTA 1292  
DB 669 GCATTATCTCGCCCAACTTCAGCTTCATGGGGAGCTGCTGCAGTTTCGAGTCCCAAGTGC 728

QY 1293 AGAACCAAGCTGGAGCATCAGGGCCAAAGAC 1324  
DB 729 TGGCCACGCTCTGTGCTGCGGAGGCTGCTAGC 760

RESULT 15  
US-09-589-184-801  
; Sequence 801, Application US/09589184  
; Patent No. 686447  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Lodes, Michael A.  
; APPLICANT: Fanger, Gary  
; APPLICANT: Vedvick, Tom  
; APPLICANT: Carter, Darrick  
; APPLICANT: Retter, Marc  
; APPLICANT: Mannion, Jane  
; APPLICANT: Retter, Marc  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND  
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.478C8  
; CURRENT APPLICATION NUMBER: US/09/589,184  
; CURRENT FILING DATE: 2000-06-05  
; NUMBER OF SEQ ID NOS: 827  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 801  
; LENGTH: 1619  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-589-184-801

Query Match 2.6%; Score 85.6; DB 4; Length 1619;  
Best Local Similarity 53.6%; Pred. No. 5.7e-16;  
Matches 178; Conservative 0; Mismatches 154; Indels 0; Gaps 0;

QY 993 ACTTTATCCCCGAGTCTCATTTCTCGGTGTGCTGTGAATGACAGCTTTTGTGAGAAA 1052  
DB 429 ACTTTGAGGACATCTACGATACAAAGTGCATCCAGTGGAGATACCAACAGCCGCA 488

QY 1053 TTTTCCGCTGTTGGCAAAATCAGTAGATTTTCATGTAGAAAGCAAAAGCTTCAATGGAT 1112  
DB 489 TCAGCTCTGTTTCATGGAAGCCATAGATACATCGATGCGGTGAAGGACTGCCGTGGGC 548

QY 1113 GTGTTCTAGTCACTGTTTACTGGATCTCCCGCTCCGCCACATCGCTATCGCCTACA 1172  
DB 549 GCGTCTGTTGATCTGCGAGGGGCACTTCGCGTGGCCACATCTGCTGGCTTACC 608

QY 1173 TCATGAAGAGATGACATCTTTTAGATGAAGCTTACAGATTTGTAAGAAAAGAC 1232  
DB 609 TGATGATGAAGAAACGGGTGAGGCTGAGGAGGCTTCGAGTTTCGTTAAGCAGCCGCA 668

QY 1233 CTACTATATCTCCAACTTCAATTTTCTGGCCAACTCTCTGGACTATGAGAAGATTA 1292  
DB 669 GCATTATCTCGCCCAACTTCAGCTTCATGGGGAGCTGCTGCAGTTTCGAGTCCCAAGTGC 728

QY 1293 AGAACCAAGCTGGAGCATCAGGGCCAAAGAC 1324  
DB 729 TGGCCACGCTCTGTGCTGCGGAGGCTGCTAGC 760

Search completed: February 28, 2004, 00:57:49  
Job time : 227 secs

Result No.	Score	Query Match	Length	DB ID	Description
1	3332	100.0	3332	9	US-09-964-277-20
2	3158	94.8	3496	9	US-09-964-277-1
3	3145	94.4	3625	12	US-10-425-114-26234
4	2350	88.5	3544	9	US-09-816-434-1
5	2950	88.5	3544	15	US-10-377-073-25
6	1795.4	53.9	1916	15	US-10-108-260A-2429
7	1762.4	52.9	2102	15	US-10-094-749-673
8	1742	52.3	2200	12	US-10-072-013-255
9	1712.4	51.4	2071	12	US-10-072-013-257
10	1560	49.8	1998	9	US-09-816-494-3
11	1560	49.8	1998	15	US-10-377-072-27
12	886.8	26.6	2558	15	US-10-404-047-1750
13	595	17.9	787	15	US-10-220-891-85
14	537.4	16.1	877	9	US-09-764-853-158
15	409.2	12.3	478	10	US-09-918-993-25801

Matches 3332: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

[illegible]

241 AAAGAATGATGTAATCGCTGGCTGCTCCAAAGCATCTTTTGTGTGGAAATGGTTATTCC 300  
Db  
241 AAAGAATGATGTAATCGCTGGCTGCTCCAAAGCATCTTTTGTGTGGAAATGGTTATTCC 300  
Qy  
301 AGTCATCTCTTTATGTAATCAAAATGTGAGGGGCTGTTTGTGGACGAGTCTCTTTGCAAGA 360  
Db  
301 AGTCATCTCTTTATGTAATCAAAATGTGAGGGGCTGTTTGTGGACGAGTCTCTTTGCAAGA 360  
Qy  
361 GCATCAACGCGGAAAGAGAAAGACATTCATCTTGGAGGGCTCTTGTGAAATGGGTT 420  
Db  
361 GCATCAACGCGGAAAGAGAAAGACATTCATCTTGGAGGGCTCTTGTGAAATGGGTT 420  
Qy  
421 TAAGTCTCTTTTGGAGGACGACACAGAGCTGACCTCATACATCTTTTGTAGTCAATGGAGT 480  
Db  
421 TAAGTCTCTTTTGGAGGACGACACAGAGCTGACCTCATACATCTTTTGTAGTCAATGGAGT 480  
Qy  
481 GGCTGAGCCTTTGAGCACACCACTTACATCATCTGTGCAAAATTAAGAAAGGAGGTGG 540  
Db  
481 GGCTGAGCCTTTGAGCACACCACTTACATCATCTGTGCAAAATTAAGAAAGGAGGTGG 540  
Qy  
541 AAAGAGGACTTATCTTGTGATGAGGACCATGAGATGATGGAACCTCAATTTGTACTGAG 600  
Db  
541 AAAGAGGACTTATCTTGTGATGAGGACCATGAGATGATGGAACCTCAATTTGTACTGAG 600  
Qy  
601 AGGTGTGCTGCTGCTGGAAGTGGAAACGGAAGAGTGTCTGCTAATTTGATAGCGGCCA 660  
Db  
601 AGGTGTGCTGCTGCTGGAAGTGGAAACGGAAGAGTGTCTGCTAATTTGATAGCGGCCA 660  
Qy  
661 TTTGTGGAATACATATCATCTCCACATTTTGGAGCCATTAATCAATGCTCCAGCTT 720  
Db  
661 TTTGTGGAATACATATCATCTCCACATTTTGGAGCCATTAATCAATGCTCCAGCTT 720  
Qy  
721 ATGAAGCGAAGTGTGAAACGAGCAAGTGTAAATACAGAGCTCATCCAGCATTCAGCG 780  
Db  
721 ATGAAGCGAAGTGTGAAACGAGCAAGTGTAAATACAGAGCTCATCCAGCATTCAGCG 780  
Qy  
781 AAACATAAGTGTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840  
Db  
781 AAACATAAGTGTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840  
Qy  
841 GATGTTGCCCTCTCTCTTCCAGACTGTTTCTCAGTCTGTTTCTCAGTCTGTTTCTCAGTCTG 900  
Db  
841 GATGTTGCCCTCTCTCTTCCAGACTGTTTCTCAGTCTGTTTCTCAGTCTGTTTCTCAGTCTG 900  
Qy  
901 AGCTTCAACTCTGTTCACTGCTTGCAGAGCTGATGACAGAGATGGATGGATGGATGGATGG 960  
Db  
901 AGCTTCAACTCTGTTCACTGCTTGCAGAGCTGATGACAGAGATGGATGGATGGATGGATGG 960  
Qy  
961 GTTAAATGCCAGCAATACCTGCTCCAGAGCTGATTTATCCCGAGTCTCATTTTCTGCG 1020  
Db  
961 GTTAAATGCCAGCAATACCTGCTCCAGAGCTGATTTATCCCGAGTCTCATTTTCTGCG 1020  
Qy  
1021 TGTGCTGTGATGACAGCTTTTGTGAGAAATTTTGTGAGAAATTTTGTGAGAAATTTTGTGAG 1080  
Db  
1021 TGTGCTGTGATGACAGCTTTTGTGAGAAATTTTGTGAGAAATTTTGTGAGAAATTTTGTGAG 1080  
Qy  
1081 TTTCAATTTGAGAAAGCAAGCTTCCATGATGATGATGATGATGATGATGATGATGATGATGAT 1140  
Db  
1081 TTTCAATTTGAGAAAGCAAGCTTCCATGATGATGATGATGATGATGATGATGATGATGATGAT 1140  
Qy  
1141 CTCCGCTCCGCAACCATCTGATCGCTACATCATGAAGAGATGGATGATGATGATGATGATGAT 1200  
Db  
1141 CTCCGCTCCGCAACCATCTGATCGCTACATCATGAAGAGATGGATGATGATGATGATGATGAT 1200  
Qy  
1201 TGAAGCTTACAGATTTGTGAGAAAGCAAGCTTCCATGATGATGATGATGATGATGATGATGAT 1260  
Db  
1201 TGAAGCTTACAGATTTGTGAGAAAGCAAGCTTCCATGATGATGATGATGATGATGATGATGAT 1260  
Qy  
1261 GGGCCAACTCTGAGCTATCAGAAAGATTTAAGAACAGATGAGATGAGATGAGATGAGATGAG 1320  
Db  
1261 GGGCCAACTCTGAGCTATCAGAAAGATTTAAGAACAGATGAGATGAGATGAGATGAGATGAG 1320

1321 GAGCAAACTCAAGCTGCTGACCTGGAGAACCAATGAACCTGCTCCTGCTGCTCCTCAGA 1380  
Db  
1321 GAGCAAACTCAAGCTGCTGACCTGGAGAACCAATGAACCTGCTCCTGCTGCTCCTCAGA 1380  
Qy  
1381 GGGTGGACAGAAAGCGAGACGCGCTCAGTCCACCTGTGCGGACTCTGCTACCTCAGA 1440  
Db  
1381 GGGTGGACAGAAAGCGAGACGCGCTCAGTCCACCTGTGCGGACTCTGCTACCTCAGA 1440  
Qy  
1441 GGCAGCAGCAAAAGGCGCGTGCATCCCGCAGGTCGCCAGCGTGCAGCGTGCAGCGTGCAGCC 1500  
Db  
1441 GGCAGCAGCAAAAGGCGCGTGCATCCCGCAGGTCGCCAGCGTGCAGCGTGCAGCGTGCAGCC 1500  
Qy  
1501 GTGCGTGTGTAGAGGACAGCGCGCTGTGTACAGGCGCTCAGTGGGCTGCACCTGTCCGAGA 1560  
Db  
1501 GTGCGTGTGTAGAGGACAGCGCGCTGTGTACAGGCGCTCAGTGGGCTGCACCTGTCCGAGA 1560  
Qy  
1561 CAGGCTGGAGACAGCAATTAAGCTCAAGCGTCTCTCTCTGATATCAAAACAGTTTC 1620  
Db  
1561 CAGGCTGGAGACAGCAATTAAGCTCAAGCGTCTCTCTCTGATATCAAAACAGTTTC 1620  
Qy  
1621 ATATTTCAGCCAGCATGGCAGCATCTTACATGGCTTCTCTCATCAGAAAGATGTTTGA 1680  
Db  
1621 ATATTTCAGCCAGCATGGCAGCATCTTACATGGCTTCTCTCATCAGAAAGATGTTTGA 1680  
Qy  
1681 ATACTACAAACCTTCCACTACTCTGTGATGGACCAACAGCTATGCCAGTTCTCCCTGT 1740  
Db  
1681 ATACTACAAACCTTCCACTACTCTGTGATGGACCAACAGCTATGCCAGTTCTCCCTGT 1740  
Qy  
1741 TCAGGAACCTATCGGAGCAGCTCCCGAAACAGTCTCTGATAAGGAGGAAGCCAGCATCCC 1800  
Db  
1741 TCAGGAACCTATCGGAGCAGCTCCCGAAACAGTCTCTGATAAGGAGGAAGCCAGCATCCC 1800  
Qy  
1801 CAAGAAGCTGACAGCGCGAGCTTTCAGACAGCAAGCGATTTGCTATTCGTCAG 1860  
Db  
1801 CAAGAAGCTGACAGCGCGAGCTTTCAGACAGCAAGCGATTTGCTATTCGTCAG 1860  
Qy  
1861 AACAGCAGCTGACAGCGCGAGCTTTCAGACAGCAAGCGATTTGCTATTCGTCAG 1920  
Db  
1861 AACAGCAGCTGACAGCGCGAGCTTTCAGACAGCAAGCGATTTGCTATTCGTCAG 1920  
Qy  
1921 CGTGGAGCAATTAACACACAGCTTCTTTTTCGCGCTTTCACACAGAGTCTTCTACTC 1980  
Db  
1921 CGTGGAGCAATTAACACACAGCTTCTTTTTCGCGCTTTCACACAGAGTCTTCTACTC 1980  
Qy  
1981 CAGCAAGTCTGCTGCTGCGCTTAAAGGCTGACCTCGGATATCTTGGCCCCCAGAC 2040  
Db  
1981 CAGCAAGTCTGCTGCTGCGCTTAAAGGCTGACCTCGGATATCTTGGCCCCCAGAC 2040  
Qy  
2041 CTCTACCCCTTCCCTGACAGCAGCTGTTATTTTGGCCACAGAGTCTTCTACTC 2100  
Db  
2041 CTCTACCCCTTCCCTGACAGCAGCTGTTATTTTGGCCACAGAGTCTTCTACTC 2100  
Qy  
2101 TGCCTCAGCCATCTAGGAGGAGTGCAGTTCCTGCTTACAGCTGCAGCGCTGCC 2160  
Db  
2101 TGCCTCAGCCATCTAGGAGGAGTGCAGTTCCTGCTTACAGCTGCAGCGCTGCC 2160  
Qy  
2161 CACTTCGAGACCAAGTCTATTCTGTGCGAGCGCGAGAGCCAAAGTGCAGAGCTGA 2220  
Db  
2161 CACTTCGAGACCAAGTCTATTCTGTGCGAGCGCGAGAGCCAAAGTGCAGAGCTGA 2220  
Qy  
2221 CTCGCGCGGAGCTGGCATGAAGAGAGCCCTTTGAAAGCAGTTTAAACGAGAGCTG 2280  
Db  
2221 CTCGCGCGGAGCTGGCATGAAGAGAGCCCTTTGAAAGCAGTTTAAACGAGAGCTG 2280  
Qy  
2281 CCAATTTGAATTTGAGAGAGCATCATGTGAGAGCAAGGTCAAGGAGAGCTGGGAA 2340  
Db  
2281 CCAATTTGAATTTGAGAGAGCATCATGTGAGAGCAAGGTCAAGGAGAGCTGGGAA 2340  
Qy  
2341 AGTGGCAGTCTAGTCTAGCTTTTTCGGGAGCATGGAATCAATGAGGCTCTCTCAGAGA 2400  
Db  
2341 AGTGGCAGTCTAGTCTAGCTTTTTCGGGAGCATGGAATCAATGAGGCTCTCTCAGAGA 2400  
Qy  
2401 AAGACACTTGTGACTCTATAGACAAATTTTTTTTCTTGTTCACAAAAAATTCCTCTGA 2460

Db	2401	AAGACAC	TGTGAC	TCTCTATAGACA	ATTTTTTTTCTTCTTACAAAAAATCCCTGGTA	2460
Qy	2461	AATCTGA	AAATATATATATG	TACATACATATATATTTTGGAAAT	TGGAGCTATGGTGTAA	2520
Db	2461	AATCTGA	AAATATATATG	TACATACATATATTTTGGAAAT	TGGAGCTATGGTGTAA	2520
Qy	2521	AAGCAAC	AGGTGGATCAAC	CCGAGTTGTCTCTTAA	CATCTGCAATTTGAGATCAGC	2580
Db	2521	AAGCAAC	AGGTGGATCAAC	CCGAGTTGTCTCTTAA	CATCTGCAATTTGAGATCAGC	2580
Qy	2581	TAAATACT	CTCTCAACAAAAT	TGAGGGGAGATGCTAGA	ATCCCCCTAGACGGAGAA	2640
Db	2581	TAAATACT	CTCTCAACAAAAT	TGAGGGGAGATGCTAGA	ATCCCCCTAGACGGAGAA	2640
Qy	2641	AACCAAT	TTTATTCAGTGA	ATTACACATCTCTTGT	TTTAAAAAGCAAGTGTCTTTGGT	2700
Db	2641	AACCAAT	TTTATTCAGTGA	ATTACACATCTCTTGT	TTTAAAAAGCAAGTGTCTTTGGT	2700
Qy	2701	GTGTGAG	GACAAATCCCT	TACCAATTTCCACG	TTGTGCTACTAGAGATCTCAAAATTT	2760
Db	2701	GTGTGAG	GACAAATCCCT	TACCAATTTCCACG	TTGTGCTACTAGAGATCTCAAAATTT	2760
Qy	2761	AGTCTTT	GTCCGGAC	CCCTTCCATAGTAC	ACCTTAGCGCTGAGACTGAGCCAGCTTGGGG	2820
Db	2761	AGTCTTT	GTCCGGAC	CCCTTCCATAGTAC	ACCTTAGCGCTGAGACTGAGCCAGCTTGGGG	2820
Qy	2821	TCAGTGA	GGTATAGAC	CCCTGTAGG	GACAGAGCTTAGTGTAAATCCAAGAGAAATGATCCT	2880
Db	2821	TCAGTGA	GGTATAGAC	CCCTGTAGG	GACAGAGCTTAGTGTAAATCCAAGAGAAATGATCCT	2880
Qy	2881	ATCCAA	GCTGATTC	CAAAACCA	CCGCTCACCTGACAGCCGAGGACACGAGCATCACTC	2940
Db	2881	ATCCAA	GCTGATTC	CAAAACCA	CCGCTCACCTGACAGCCGAGGACACGAGCATCACTC	2940
Qy	2941	TGCTGGA	CGGACCAAT	TAGGGG	CCCTTGCCAAAGTCTACCTTAGAGCAAAACCAAGTACCTCA	3000
Db	2941	TGCTGGA	CGGACCAAT	TAGGGG	CCCTTGCCAAAGTCTACCTTAGAGCAAAACCAAGTACCTCA	3000
Qy	3001	GACAGGA	AGTTCGGG	CTTTGA	CCATACATATCTGGTAGCCCATTTTCTTAGGCATTTGT	3060
Db	3001	GACAGGA	AGTTCGGG	CTTTGA	CCATACATATCTGGTAGCCCATTTTCTTAGGCATTTGT	3060
Qy	3061	GAATAGG	TAGTACTG	TACATCTTT	TGAGCCCAATTCAAACTCTCTATGSCACAAAAT	3120
Db	3061	GAATAGG	TAGTACTG	TACATCTTT	TGAGCCCAATTCAAACTCTCTATGSCACAAAAT	3120
Qy	3121	CCCGTGG	CCCTAGAT	TGGAGATA	TAATTTTTTTTCTCTCAGCTTTATGAAGAGAGGGAAA	3180
Db	3121	CCCGTGG	CCCTAGAT	TGGAGATA	TAATTTTTTTTCTCTCAGCTTTATGAAGAGAGGGAAA	3180
Qy	3181	CHGTCT	AGGATTCAGCTGA	ACCACGAG	AACTGGGCAACATCACGATTTAGCTTAAGGTT	3240
Db	3181	CHGTCT	AGGATTCAGCTGA	ACCACGAG	AACTGGGCAACATCACGATTTAGCTTAAGGTT	3240
Qy	3241	GGGAGG	CTAA	CGAGTCTACCT	CCCTCTTCTTAATCAAAGAAATCTGTTTAAATGGGATG	3300
Db	3241	GGGAGG	CTAA	CGAGTCTACCT	CCCTCTTCTTAATCAAAGAAATCTGTTTAAATGGGATG	3300
Qy	3301	TCAATC	CTTTAAATAA	AGATGA	ACTTGGTTTC	3332
Db	3301	TCAATC	CTTTAAATAA	AGATGA	ACTTGGTTTC	3332

## RESULT 2

```

RECORD 2
US-09-964-277-1
; Sequence 1, Application US/09964277
; Patent No. US20020137170A1
; GENERAL INFORMATION:
; APPLICANT: Luche, Ralf M.
; APPLICANT: Wei, Bo
; TITLE OF INVENTION: DSP-16 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125.434

```

901 AGCTTCAACTCTGTTTCACTGCTTGC----- 926  
901 AGCTTCAACTCTGTTTCACTGCTTGCAGGTGGTGTGCTGAGTTCTCTCGTTGTTCCCT 960  
927 ----- 926  
961 GGCTCTGTGAAGAAATCCACTAGTCCCTACCTGCAATTTCTCAGCCTTGTCTACCT 1020  
927 ----- 926  
1021 GTTGCACAATTTGGGCAACCCGAAATTTCTCCAAATCTTTATCTTGGTGCAGGAGAT 1080  
927 ----- AGAGCTGATGACAGAGATGGGATGTTGTTATGTTTAAATGCCAGCAAT 976  
1081 GTCTCTCAACAGGAGCTGATGACAGAGATGGGATGTTGTTATGTTTAAATGCCAGCAAT 1140  
977 ACCTGTCCAAAGCTGACTTTTATCCCGAGTCTCATTTCTCGTGTGCTGTGAATGAC 1036  
1141 ACCTGTCCAAAGCTGACTTTTATCCCGAGTCTCATTTCTCGTGTGCTGTGAATGAC 1200  
1037 AGCTTTTGTGAGAAATTTTGCCTGTTGGACAAATCAGTAGATTTTCAATGAGAAAGCA 1096  
1201 AGCTTTTGTGAGAAATTTTGCCTGTTGGACAAATCAGTAGATTTTCAATGAGAAAGCA 1260  
1097 AAAGCCTCCAATGGATGTTCTTAGTGCACTGTTTAGCTGGATCTCCCGTCCGCCACC 1156  
1261 AAAGCCTCCAATGGATGTTCTTAGTGCACTGTTTAGCTGGATCTCCCGTCCGCCACC 1320  
1157 ATCGCTATCGCTTACATCATGAGAGGATGGACATGTTCTTTAGATGAGCTTACAGATTT 1216  
1321 ATCGCTATCGCTTACATCATGAGAGGATGGACATGTTCTTTAGATGAGCTTACAGATTT 1380  
1217 GTGAAAGAAAAAGACTACTATATCTCCAAATTTCAATTTTCTGGGCCAACTCCTGGAC 1276  
1381 GTGAAAGAAAAAGACTACTATATCTCCAAATTTCAATTTTCTGGGCCAACTCCTGGAC 1440  
1277 TATGAGAAAGAAATTAAGAACACAGACTGAGGATCAGGGCCAAAGAGCAAACTCAAGCTG 1336  
1441 TATGAGAAAGAAATTAAGAACACAGACTGAGGATCAGGGCCAAAGAGCAAACTCAAGCTG 1500  
1337 CTGCACTCGAGAAAGCAATGAACCTGTCCCTGCTGTCTCAGAGGGTGGACAGAAAGC 1396  
1501 CTGCACTCGAGAAAGCAATGAACCTGTCCCTGCTGTCTCAGAGGGTGGACAGAAAGC 1560  
1397 GAGACGCCCTCAGTCCACCTGTGCGGACTGTGTAACCTCAGAGGCGACAGGACAAAGG 1456  
1561 GAGACGCCCTCAGTCCACCTGTGCGGACTGTGTAACCTCAGAGGCGACAGGACAAAGG 1620  
1457 CCGTGCATCCCGCAGCGTCCCGCTGCGGCTGCGGCTGCGGCTGCTGTAGAGGAC 1516  
1621 CCGTGCATCCCGCAGCGTCCCGCTGCGGCTGCGGCTGCGGCTGCTGTAGAGGAC 1680  
1517 AGCCCGCTGTGACAGCGCTCAGTGGGCTGACCTGTGCGCAGACAGGCTGGAAGACAGC 1576  
1681 AGCCCGCTGTGACAGCGCTCAGTGGGCTGACCTGTGCGCAGACAGGCTGGAAGACAGC 1740  
1577 AATAAGCTCAAGCGTTCCTTCTCTGAGATCAAAATCAGTTTCAATTTAGCCAGCATG 1636  
1741 AATAAGCTCAAGCGTTCCTTCTCTGAGATCAAAATCAGTTTCAATTTAGCCAGCATG 1800  
1637 GCAGCATCCTTACATGGCTTCTCCTCATCAGAAAGATGCTTTGGAATACCTACAAACCTTCC 1696  
1801 GCAGCATCCTTACATGGCTTCTCCTCATCAGAAAGATGCTTTGGAATACCTACAAACCTTCC 1860  
1697 ACTACTCTGATGGGACCAAAAGCTATGCCAGTTCTCCCTGTGTTTCAAGAACTATCGGAG 1756  
1861 ACTACTCTGATGGGACCAAAAGCTATGCCAGTTCTCCCTGTGTTTCAAGAACTATCGGAG 1920  
1757 CAGACTCCGAAACCAAGTCTCTGATAGGAGGAGGACGACATCCCAAGAGCTGACAGC 1816  
1921 CAGACTCCGAAACCAAGTCTCTGATAGGAGGAGGACGACATCCCAAGAGCTGACAGC 1980

1817 GCCAGGCTTTCAGACAGCCAGCAAGGATTCATTGCTCGGTCAAGAACACAGCAGCTGGC 1876  
1981 GCCAGGCTTTCAGACAGCCAGCAAGGATTCATTGCTCGGTCAAGAACACAGCAGCTGGC 2040  
1877 ACCGCCCTTAAGGGCTGGCACTCGGATATCTTGGCCCTCCAGAGCTTACCCCTTCCCTG 2056  
2041 ACCGCCCTTAAGGGCTGGCACTCGGATATCTTGGCCCTCCAGAGCTTACCCCTTCCCTG 2220  
1937 CACACAGCTTTCCTTTTCGGCTTTCACACAGCAGCAGCAGCTCAGAACTGCTGCTGGC 1996  
2101 CACACAGCTTTCCTTTTCGGCTTTCACACAGCAGCAGCAGCTCAGAACTGCTGCTGGC 2160  
1997 CTGGCCCTTAAGGGCTGGCACTCGGATATCTTGGCCCTCCAGAGCTTACCCCTTCCCTG 2056  
2161 CTGGCCCTTAAGGGCTGGCACTCGGATATCTTGGCCCTCCAGAGCTTACCCCTTCCCTG 2220  
2057 ACCAGCAGCTGGTATTTTGGCCAGAGTCTCAGCTTCTACTCTGCTCAGCACTAC 2116  
2221 ACCAGCAGCTGGTATTTTGGCCAGAGTCTCAGCTTCTACTCTGCTCAGCACTAC 2280  
2117 GGAGCAGTGGCAGTTACTCTGCTTACAGCTGACAGCTGCGCCTTGGGAGACCA 2176  
2281 GGAGCAGTGGCAGTTACTCTGCTTACAGCTGACAGCTGCGCCTTGGGAGACCA 2340  
2177 GTCTATCTGTGCGCAGCGCCAGCAAGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2356  
2341 GTCTATCTGTGCGCAGCGCCAGCAAGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2400  
2237 CATGAAGAGAGCCCTTTGAAAGCAGTTTAAACGAGAGCTGCGCAATTTGGA 2296  
2401 CATGAAGAGAGCCCTTTGAAAGCAGTTTAAACGAGAGCTGCGCAATTTGGA 2460  
2297 GAGACATCATGTGAGAAACAGCTGACGGAAGCTGCGGAAAGTGGGAGTCACTCT 2356  
2461 GAGACATCATGTGAGAAACAGCTGACGGAAGCTGCGGAAAGTGGGAGTCACTCT 2520  
2357 AGCTTTTCGGCAGCATGGAATCATTTAGCTGCTCCTGAGAAAGACACACTTGTGACTT 2416  
2521 AGCTTTTCGGCAGCATGGAATCATTTAGCTGCTCCTGAGAAAGACACACTTGTGACTT 2580  
2417 CTATAGACAAATTTTCTTGTTCACAAAAAATTTCCCTGTAATCTGAAATATATAT 2476  
2581 CTATAGACAAATTTTCTTGTTCACAAAAAATTTCCCTGTAATCTGAAATATATAT 2640  
2477 ATGTACATACATATATTTTGGAAATGAGCTATGTTGTAAGCAACAGCTGATC 2536  
2641 ATGTACATACATATATTTTGGAAATGAGCTATGTTGTAAGCAACAGCTGATC 2700  
2537 AACCCAGTTGTTACTCTCTTAACATCTGCAATTTGAGAGATCAGCTAATCTCTCTCAAC 2596  
2701 AACCCAGTTGTTACTCTCTTAACATCTGCAATTTGAGAGATCAGCTAATCTCTCTCAAC 2760  
2597 AAAAATGGAAGGCGAGATGCTAGAAATCCCTTACAGCGAGGAAACCAATTTTATCAGT 2656  
2761 AAAAATGGAAGGCGAGATGCTAGAAATCCCTTACAGCGAGGAAACCAATTTTATCAGT 2820  
2657 GAATTTACATCTCTCTTCTTAAAGCAAGTGTCTTTGTTGTTGGAGGACAAATC 2716  
2821 GAATTTACATCTCTCTTCTTAAAGCAAGTGTCTTTGTTGTTGGAGGACAAATC 2880  
2717 CCGTACCAATTTTCCAGTGTGCTACTAAGAGATCTCAAAATATTAGTCTTTGTCGGACC 2776  
2881 CCGTACCAATTTTCCAGTGTGCTACTAAGAGATCTCAAAATATTAGTCTTTGTCGGACC 2940  
2777 CTTTCATAGTACACCTTAGCTGAGCTGAGCTGAGCTGAGGCTGAGGCTGAGTAGACCC 2836  
2941 CTTTCATAGTACACCTTAGCTGAGCTGAGCTGAGCTGAGGCTGAGGCTGAGTAGACCC 3000  
2837 TGTTAGGACACAGAGCTAGTGGTAAATCCAGAGAAATGATCTCTATCCAAAGCTGATTCA 2896  
3001 TGTTAGGACACAGAGCTAGTGGTAAATCCAGAGAAATGATCTCTATCCAAAGCTGATTCA 3060  
2897 CAAACCCAGGCTCAGCTGACAGCCGAGGACACAGAGCTCACTCTGCTGACGACCAATT 2956





1392	Db		AAGCCTCCAAATGGAGATGGTTCTCTAGTGCACCTGTTTACGTGGGATCTCCGGCTCCGCCACCA	1451
1158	QY	TCGCTATCGCCTCATCATGAAGAGGATGGACATGTCCTTTAGATGAAGCTTACAGATTG	1217	
1452	Db	TCGCTATCGCCTCATCATGAGAGGANTGGACNTGCTTTAGATGAAGCTTACAGATTG	1511	
1218	QY	TGAAGAAAAAGACCTCATATATCTCCAAACTTCAATTTTCTGGGCCAACTCTCGACT	1277	
1512	Db	TGAAGAAAAAGACCTACTATATCTCCAACTTCAATTTTCTGGGCCAACTCTCGACT	1571	
1278	QY	ATGAGAAAGACATTAAAGAACAGACATGGAGCATCAGGGGCCAAAGAGCAAACTCAAGCTGC	1337	
1572	Db	ATGAGAAAGAGTTTAAAGAACAGACTTGGAGCATCAGGGCCAAAGAGCAAACTCAAGCTGC	1631	
1338	QY	TGCACCTGGAGAACCCAAATGAACCTGTCCCTGCTGTCTCAGAGGGTGGACAGAAAAAGCG	1397	
1632	Db	TGCACCTGGAGAACCCAAATGAACCTGTCCCTGCTGTCTCAGAGGGTGGACAGAAAAAGCG	1691	
1398	QY	AGAGGCCCTCAGTCGACCCCTGTGGCGATCTCTGCTACCTCAGAGCGACGAGCAAAAGGC	1457	
1692	Db	AGAGGCCCTCAGTCCACCCCTGTGGCGATCTCTGCTACCTCAGAGCGACGAGCAAAAGGC	1751	
1458	QY	CCGTGCATCCCGCCAGAGCGTGCCACAGCGTGCCAGCGCTGCAGCGCTCGCTGTTAGAGACA	1517	
1752	Db	CCGTGCATCCCGCCAGAGCGTGCCACAGCGTGCCACAGCGTGCGCTGTTAGAGACA	1811	
1518	QY	GCCCGCTGTTACAGGCGGTGAGTGGGTGCACTGTCGCCAGACAGGCTGGAAGACAGCA	1577	
1812	Db	GCCCGCTGTTACAGGCGGTGAGTGGGTGCACTGTCGCCAGACAGGCTGGAAGACAGCA	1871	
1578	QY	ATAAGCTCAAGCGTTCTTCTCTCTGGATCAAAATCAGTTTCAATTCAGCCAGCATGG	1637	
1872	Db	ATAAGCTCAAGCGTTCTTCTCTCTGGATCAAAATCAGTTTCAATTCAGCCAGCATGG	1931	
1638	QY	CAGCATCCTTACATGGCTTCTCTCATCAGAGAGATGCTTTGGAATCATCAAACTTTCCA	1697	
1932	Db	CAGCATCCTTACATGGCTTCTCTCATCAGAGAGATGCTTTGGAATCATCAAACTTTCCA	1991	
1698	QY	CTACTCTGGATGGGAGCAACAAAGCTATGCCAGTTCTCCCTGTTCCAGAACTATCGGAGC	1757	
1992	Db	CTACTCTGGATGGGAGCAACAAAGCTATGCCAGTTCTCCCTGTTCCAGAACTATCGGAGC	2051	
1758	QY	AGACTCCCGAAACCAAGTCTCTGATAGGAGGAAGCCAGCATCCCCAGAGCTGCAGACCG	1817	
2052	Db	AGACTCCCGAAACCAAGTCTCTGATAGGAGGAAGCCAGCATCCCCAGAGCTGCAGACCG	2111	
1818	QY	CAGAGCCTTCAGACAGCCAGAGCAAGCGATTGCAATTCGGTTCAGAACAGAGCATGGCA	1877	
2112	Db	CAGAGCCTTCAGACAGCCAGAGCAAGCGATTGCAATTCGGTTCAGAACAGAGCATGGCA	2171	
1878	QY	CCGCCACAGAGTGCCCTTTTATCTCCAATGCATCGAAAGTGGAGCGGTGAGAGCAATTACC	1937	
2172	Db	CCGCCACAGAGTGCCCTTTTATCTCCAATGCATCGAAAGTGGAGCGGTGAGAGCAATTACC	2231	
1938	QY	AACACAGCTTGCTTTTTCGGCCTTTTCCACAGCCAGCAGCACTTCACGAAAGTCTGCTGCC	1997	
2232	Db	AACACAGCTTGCTTTTTCGGCCTTTTCCACAGCCAGCAGCACTTCACGAAAGTCTGCTGCC	2291	
1998	QY	TGGGCTTTAAGGGCTGGCACCTCGGATATCTTGGCCCCCAGACCTCTACCCCTTCCCTGA	2057	
2292	Db	TGGGCTTTAAGGGCTGGCACCTCGGATATCTTGGCCCCCAGACCTCTACCCCTTCCCTGA	2351	
2058	QY	CCAGCAGCTGGTATTTTGGCCACAGAGTCCTCACACTTCTACTCTGCCTCAGGCCATCTACG	2117	
2352	Db	CCAGCAGCTGGTATTTTGGCCACAGAGTCCTCACACTTCTACTCTGCCTCAGGCCATCTACG	2411	
2118	QY	GAGGCAGTGCAGTTACTCTGCTTACAGCTGCGAGCCAGCTGCCCCATTTGCGGAGCCAG	2177	
2412	Db	GAGGCAGTGCAGTTACTCTGCTTACAGCTGCGAGCCAGCTGCCCCATTTGCGGAGCCAG	2471	
2178	QY	TCTATTCTGTGCGAGGCGGACAGAACGCAAGTCAGACAGCTGACTCGCGGGGAGACTGGC	2237	

2472	TCTATTCTGTGCGGAGCGGGGAGAGCCAAAGTAGCAGAGCTGACTGCGCGCGGAGCTGGC	2531
2238	ATGAAGAGAGCCCTTTGAAAGCAGGTTTAAACGCGAAGAGCTGCCAAATGGAAATTTGGAG	2297
2532	ATGAAGAGAGCCCTTTGAAAGCAGGTTTAAACGCGAAGAGCTGCCAAATGGAAATTTGGAG	2591
2298	AGAGCATCATGTCCAGAGAACAGGTCCAGGGAGAGCTGGGGAAAGTGGCGAGTCAGTCTA	2357
2592	AGAGCATCATGTCCAGAGAACAGGTCCAGGGAGAGCTGGGGAAAGTGGCGAGTCAGTCTA	2651
2358	GCTTTTTCGGGAGCATGGAATCATTTGAGGTCTCTGAGAGAGAAAGACACTTGTGACTTC	2417
2652	GCTTTTTCGGGAGCATGGAATCATTTGAGGTCTCTGAGAGAGAAAGACACTTGTGACTTC	2711
2418	TATAGACAAATTTTTTTTCTTGTTCACAAAAAATTCCTGTAAATCTGAAATATATATA	2477
2712	TATAGACAAATTTTTTTTCTTGTTCACAAAAAATTCCTGTAAATCTGAAATATATATA	2771
2478	TGTACATACATATATATTTTTTGGAAAAATGGAGCTATGTGTAAAAAGCAACAGGTGGATCA	2537
2772	TGTACATACATATATATTTTTTGGAAAAATGGAGCTATGTGTAAAAAGCAACAGGTGGATCA	2831
2538	ACCAGTTGTTACTCTCTTAACATCTGCATTTGAGAGATCAGCTAATACTTCTTCAACA	2597
2832	ACCAGTTGTTACTCTCTTAACATCTGCATTTGAGAGATCAGCTAATACTTCTTCAACA	2891
2598	AAAAATGGAGGGCAGATGCTAGAAATCCCCCTAGACGGAGGAAAAACCTATTTATTCAGTG	2657
2892	AAAAATGGAGGGCAGATGCTAGAAATCCCCCTAGACGGAGGAAAAACCTATTTATTCAGTG	2951
2658	AATTACATACCTCTTGTCTTTAAAAAGCAAGTGTCTTTGTGTGGAGGACAAAAATCC	2717
2952	AATTACATACCTCTTGTCTTTAAAAAGCAAGTGTCTTTGTGTGGAGGACAAAAATCC	3011
2718	CCTACCATTTTCCAGTTGTGCTACTAAGAGATCTCAATATATAGTCTTTGTCGGGACCC	2777
3012	CCTACCATTTT--CAGCTGTGTGCTACTAAGAGATCTCAATATATAGTCTTTGTCGGGACCC	3070
2778	TTTCCATAGTACACCTTAGCGCTGAGACTGAGCCAGCTTCGGGGCTCAGGTAGGTAGACCCCT	2837
3071	TTCCATAGTACACTTAGCGCTGAGACTGAGCCAGCTTCGGGGCTCAGGTAGGTAGACCCCT	3130
2838	GTTAGGGACAGAGCCTAGTGTAAATCCAGAGAGAAATGATCTCTATCCAAAGCTGATTCAC	2897
3131	GTTAGGGACAGAGCCTAGTGTAAATCCAGAGAGAAATGATCTCTATCCAAAGCTGATTCAC	3190
2898	AAACCCACGCTCACCTGACGCGAGGAGACAGAGCATCACTCTGCTGACGGAGCAATTA	2955
3191	AAACCCACGCTCACCTGACGCGAGGAGACAGAGCATCACTCTGCTGACGGAGCAATTA	3250
2958	GGGGCCTTGCCAGAGTCTACCTTAGACCAACCCAGTACTCCAGACAGAGAAAGTCGGGGC	3017
3251	GGGGCCTTGCCAGAGTCTACCTTAGACCAACCCAGTACTCCAGACAGAGAAAGTCGGGGC	3310
3018	TTTAGCCACTACCATATCTGTGTAGCCCAATTTTCTAGGCATTTGTGAATAGGTAGGTAGCTA	3077
3311	TTTAGCCACTACCATATCTGTGTAGCCCAATTTTCTAGGCATTTGTGAATAGGTAGGTAGCTA	3370
3078	GTACACACTTTTTCAGACCAATTTCAAACCTGTCTATGCACAAAAATTCOCGTGGGCTAGATGG	3133
3371	GTACACACTTTTTCAGACCAATTTCAAACCTGTCTATGCACAAAAATTCOCGTGGGCTAGATGG	3430
3138	AGATAATTTTTTTTTTCTCTCAGCTTTATGAGAGAGAGGAAACTGTCTAGGAATTCAGCT	3197
3431	AGATAATTTTTTTTTTCTCTCAGCTTTATGAGAGAGAGGAAACTGTCTAGGAATTCAGCT	3490
3198	GAACACACAGGAACCTGGCAACATCAACGATTTAAGCTTTGGAGGCTTAACGAGTCT	3257
3491	GAACACACAGGAACCTGGCAACATCAACGATTTAAGCTTTGGAGGCTTAACGAGTCT	3550
3258	ACCTCCCTCTTTGTAATAACAAAGAAATTTGTTTAAAAATGGGATTTGTCAATCCTTTAAATAAA	3317
3551	ACCTCCCTCTTTGTAATAACAAAGAAATTTGTTTAAAAATGGGATTTGTCAATCCTTTAAATAAA	3610

QY 3318 GATGAACCTGGTTTC 3332  
Db 3611 GATGAACCTGGTTTC 3625

## RESULT 4

US-09-816-494-1  
; Sequence 1, Application US/09816494  
; Patent No. US20020034807A1  
; GENERAL INFORMATION:  
; APPLICANT: Meyers, Rachel A.  
; TITLE OF INVENTION: 38692 AND 21117, NOVEL DUAL SPECIFICITY  
; FILE REFERENCE: 10448-030002  
; CURRENT APPLICATION NUMBER: US/09/816,494  
; CURRENT FILING DATE: 2001-03-23  
; PRIOR APPLICATION NUMBER: US 60/191,858  
; PRIOR FILING DATE: 2000-03-24  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 3544  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (589)...(2583)  
US-09-816-494-1

Query Match 88.5%; Score 2950; DB 9; Length 3544;  
Best Local Similarity 95.0%; Pred. No. 0;  
Matches 3135; Conservative 0; Mismatches 0; Indels 165; Gaps 2;  
QY 197 GCTTTCAGTCCAGTGTAAGCTGTGGAGCGCGGAGCAAGGTAAGAATGATGTAATG 256  
Db 224 GCTTTCAGTCCAGTGTAAGCTGTGGAGCGCGGAGCAAGGTAAGAATGATGTAATG 283  
QY 257 CGCTGGCTGCTCCAAAGCATCTTTGTTGGAAATGTTATTCAGTCATCTCTTTATGA 316  
Db 284 CGCTGGCTGCTCCAAAGCATCTTTGTTGGAAATGTTATTCAGTCATCTCTTTATGA 343  
QY 317 ATCAAAATGAGGGGCTGCTTTGTTGGACGGAGTCCTTTGCAAGACATCAACGGGAAA 376  
Db 344 ATCAAAATGAGGGGCTGCTTTGTTGGACGGAGTCCTTTGCAAGACATCAACGGGAAA 403  
QY 377 GAGAAAGAGACATTCACCTGGAGGGCTCTTGCTGAAATGGTTTAACCTCCTTTTGGC 436  
Db 404 GAGAAAGAGACATTCACCTGGAGGGCTCTTGCTGAAATGGTTTAACCTCCTTTTGGC 463  
QY 437 AGTCACCAACCGCTGACCTCATACATCTTTTAGTACAAATGGAGTGGCTGAGCCTTTGAGC 496  
Db 464 AGTCACCAACCGCTGACCTCATACATCTTTTAGTACAAATGGAGTGGCTGAGCCTTTGAGC 523  
QY 497 ACACCAACCATTCATCATCTGTCGCAATTAAGNAGAGGTGGGAAAGAGGACTTATTG 556  
Db 524 ACACCAACCATTCATCATCTGTCGCAATTAAGNAGAGGTGGGAAAGAGGACTTATTG 583  
QY 557 TTGTCATGGCCCATGAGATGATTGGAATCAAAATGTTTACTGAGAGGTGGTGGCTCTGC 616  
Db 584 TTGTCATGGCCCATGAGATGATTGGAATCAAAATGTTTACTGAGAGGTGGTGGCTCTGC 643  
QY 617 TGGAAAGTGGAAACGGAAAGGAGTGTCTAATAGNAGCGGCAATTTGTTGGAATACAATA 676  
Db 644 TGGAAAGTGGAAACGGAAAGGAGTGTCTAATAGNAGCGGCAATTTGTTGGAATACAATA 703  
QY 677 CATCCCAATTTTGAAGCCCAATTAATCAACTGCTCCAAAGCTTATGAAGCGAAGGTTC 736  
Db 704 CATCCCAATTTTGAAGCCCAATTAATCAACTGCTCCAAAGCTTATGAAGCGAAGGTTC 763  
QY 737 AACAGGACAAAGTGTATTAACAGAGCTCATCCAGCATTCAGGAAACATAAGGTTGACA 796  
Db 764 AACAGGACAAAGTGTATTAACAGAGCTCATCCAGCATTCAGGAAACATAAGGTTGACA 823

QY 797 TTGATTGCAGTCAGAGGTTGTAGTTTACGATCAAAGCTCCCAAGATGTTGCTCTCTCT 856  
Db 824 TTGATTGCAGTCAGAGGTTGTAGTTTACGATCAAAGCTCCCAAGATGTTGCTCTCTCT 883  
QY 857 CTTTCAGACTGTTTCTCAGCTACTTCTGGTAACTGGAGAGAGCTTCAACTCTGTTTC 916  
Db 884 CTTTCAGACTGTTTCTCAGCTACTTCTGGTAACTGGAGAGAGCTTCAACTCTGTTTC 943  
QY 917 ACCTGCTTGC----- 926  
Db 944 ACCTGCTTGCAGGTGGGTTTGCTGAGTTCTCTCGTTGTTTCCCTGGCCCTCTGTGAAGAA 1003  
QY 927 ----- 926  
Db 1004 AATCCACTCTAGTCCCTACCTGCATTTCTCAGCCTTGTCTTACCTGTTGCCAATTTGGGC 1063  
QY 927 -----AGGAGC 932  
Db 1064 CAACCCGAATCTTCCCAATCTTTATCTTTGGCTGCCAGAGATGTCCTCAACAGGAGC 1123  
QY 933 TGATGAGAGCAATGGGATTTGTTTAAATGCCAGCAATACCTGTCCAAAGCCTTG 992  
Db 1124 TGATGAGAGCAATGGGATTTGTTTAAATGCCAGCAATACCTGTCCAAAGCCTTG 1183  
QY 993 ACTTTATCCCGAGTCTCAATTTCCGCTGCTGCTGAATGACAGCTTTGTGAGAAA 1052  
Db 1184 ACTTTATCCCGAGTCTCAATTTCCGCTGCTGCTGAATGACAGCTTTGTGAGAAA 1243  
QY 1053 TTTTGGCGTGTGGACAAATCAGTAGATTTTCAATGAGAAAGCAAAAGCCTCCAATGGAT 1112  
Db 1244 TTTTGGCGTGTGGACAAATCAGTAGATTTTCAATGAGAAAGCAAAAGCCTCCAATGGAT 1303  
QY 1113 GTGTTCTAGTGCATGTTTAGCTGGGATCTCCGCTCCGCCACCATCGCTATGCGCTACA 1172  
Db 1304 GTGTTCTAGTGCATGTTTAGCTGGGATCTCCGCTCCGCCACCATCGCTATGCGCTACA 1363  
QY 1173 TCATGAAGAGGATGGACATGTCCTTTAGATGAAGCTTACAGATTTGTGAAGAAAAAGAC 1232  
Db 1364 TCATGAAGAGGATGGACATGTCCTTTAGATGAAGCTTACAGATTTGTGAAGAAAAAGAC 1423  
QY 1233 CTACTATATCTCAAATCTCAATTTTCTGGGCCAATCTCTGGAATATGAGAAGAGATTA 1292  
Db 1424 CTACTATATCTCAAATCTCAATTTTCTGGGCCAATCTCTGGAATATGAGAAGAGATTA 1483  
QY 1293 AGAACAGACTGGAGCATCAGGCCCAAGAGCAAACTCAAGTGTCTGCACCTGGGAGAGC 1352  
Db 1484 AGAACAGACTGGAGCATCAGGCCCAAGAGCAAACTCAAGTGTCTGCACCTGGGAGAGC 1543  
QY 1353 CAAATGAACCTGTCCTGCTCTCAGAGGGTGGACAGAAAAAGCGAGACGCCCTCAGTC 1412  
Db 1544 CAAATGAACCTGTCCTGCTCTCAGAGGGTGGACAGAAAAAGCGAGACGCCCTCAGTC 1603  
QY 1413 CACCTGTGCGGACTGCTACCTCAGAGGCGAGGACAAAGGCCCGTGCATCCGCCCA 1472  
Db 1604 CACCTGTGCGGACTGCTACCTCAGAGGCGAGGACAAAGGCCCGTGCATCCGCCCA 1663  
QY 1473 GCGTGCCCAAGTGGCCAGCGTGCAGCGCTCGCTGTGTAGAGGACAGCCCGCTGGTACAGG 1532  
Db 1664 GCGTGCCCAAGTGGCCAGCGTGCAGCGCTCGCTGTGTAGAGGACAGCCCGCTGGTACAGG 1723  
QY 1533 CGCTCAGTGGGCTGACCTGTCCGAGACAGGCTGGAAGACAGCAATAGCTCAAGCGTT 1592  
Db 1724 CGCTCAGTGGGCTGACCTGTCCGAGACAGGCTGGAAGACAGCAATAGCTCAAGCGTT 1783  
QY 1593 CCTTCTCTCTGGATATCAAAATCAGTTTCTATATTCAGCCAGCATGGCAGCATCTTACATG 1652  
Db 1784 CCTTCTCTCTGGATATCAAAATCAGTTTCTATATTCAGCCAGCATGGCAGCATCTTACATG 1843  
QY 1653 GCTTCTCTCTCATCAGAAAGTGTCTTGGAAATCTATAAACTCTTCCACTACTCTGGAGTGGAA 1712  
Db 1844 GCTTCTCTCTCATCAGAAAGTGTCTTGGAAATCTATAAACTCTTCCACTACTCTGGAGTGGAA 1903

Qy	1713	CCAA	CAAGCTATGCGAGTCTCCCTCTTTCAGGAACTATCGGAGCAGACTCCCGAAACCA	1772
Db	1904	CCAA	CAAGCTATGCGAGTCTCCCTCTTTCAGGAACTATCGGAGCAGACTCCCGAAACCA	1963
Qy	1773	GTCT	CTGATAAGGAGGAGCCAGCATCCCAGAAGCTGCAGACCGCAGGCGTTCAGACA	1832
Db	1964	GTCT	CTGATAAGGAGGAGCCAGCATCCCAGAAGCTGCAGACCGCAGGCGTTCAGACA	2023
Qy	1833	GCC	GAGCAAGCGATTGCAATTCGTTCAGAACCAACGACGAGTGGCACCGCCAGAGGTCCC	1892
Db	2024	GCC	GAGCAAGCGATTGCAATTCGTTCAGAACCAACGACGAGTGGCACCGCCAGAGGTCCC	2083
Qy	1893	TTTT	TATCTCCATCGATCGAAGTGGGAGCGTGGAGGACAAATTACCACACCAAGCTTCCTTT	1952
Db	2084	TTTT	TATCTCCATCGATCGAAGTGGGAGCGTGGAGGACAAATTACCACACCAAGCTTCCTTT	2143
Qy	1953	TCGG	CTTTTCCACGACCCAGCAGCACCTCCACGAAGTCTCTCGCCCTGGGCGCTTAAGGGCT	2012
Db	2144	TCGG	CTTTTCCACGACCCAGCAGCACCTCCACGAAGTCTCTCGCCCTGGGCGCTTAAGGGCT	2203
Qy	2013	GGCA	CTCGGATCTTGGCCCCCAGACCTCTACCCCTTCCCTGACACGACGCTGGTATT	2072
Db	2204	GGCA	CTCGGATCTTGGCCCCCAGACCTCTACCCCTTCCCTGACACGACGCTGGTATT	2263
Qy	2073	TTTG	CCACAGAGTCTCTACACTTCTACTCTGCCCTCAGCCCATCTACGGAGCGACAGTTC	2132
Db	2264	TTTG	CCACAGAGTCTCTACACTTCTACTCTGCCCTCAGCCCATCTACGGAGCGACAGTTC	2323
Qy	2133	ACT	CTGCCTACAGCTGACGAGCTGCCCATCTTTCGGGAGCAACAGTCTATTCTGTGGCGA	2192
Db	2324	ACT	CTGCCTACAGCTGACGAGCTGCCCATCTTTCGGGAGCAACAGTCTATTCTGTGGCGA	2383
Qy	2193	GGCG	CAGAAGTGACACAGACTCAGCTCGCGCGGAGCTGGCATGAAGAGAGCCCT	2252
Db	2384	GGCG	CAGAAGTGACACAGACTCAGCTCGCGCGGAGCTGGCATGAAGAGAGCCCT	2443
Qy	2253	TTG	AAAAAGCGTTTAAACGCGAAGTGCCTCAATTGGAAATTTGGAGAGAGCATCATGTGAG	2312
Db	2444	TTG	AAAAAGCGTTTAAACGCGAAGTGCCTCAATTGGAAATTTGGAGAGAGCATCATGTGAG	2503
Qy	2313	AGAA	CAGGCTCAGGAGAGAGCTGGGAAAGTGGGAGTCAGTCTAGCTTTTCGGGCGAGCA	2372
Db	2504	AGAA	CAGGCTCAGGAGAGAGCTGGGAAAGTGGGAGTCAGTCTAGCTTTTCGGGCGAGCA	2563
Qy	2373	TGGA	AAATCATTTGAGGTCTCTTGAGAAGAAAGACACTTGTGACTTCTATAGACAAATTTTTT	2432
Db	2564	TGGA	AAATCATTTGAGGTCTCTTGAGAAGAAAGACACTTGTGACTTCTATAGACAAATTTTTT	2623
Qy	2433	TTT	CTTGTTCACAAAAAATTTCCCTGTAAATCTGAAATATNTATATGTACATACATATAT	2492
Db	2624	TTT	CTTGTTCACAAAAAATTTCCCTGTAAATCTGAAATATNTATATGTACATACATATAT	2683
Qy	2493	ATTT	TTGAAAAATGGAGCTATGGTGTAAAAAGCAACAGGTGGATCAACCCAGTTGTACTCTC	2552
Db	2684	ATTT	TTGAAAAATGGAGCTATGGTGTAAAAAGCAACAGGTGGATCAACCCAGTTGTACTCTC	2743
Qy	2553	TCT	TAACTATCTGCATTTTGAGAGATCAAGCTAATATCTCTCTCAACAAAAATGGAGGGGAG	2612
Db	2744	TCT	TAACTATCTGCATTTTGAGAGATCAAGCTAATATCTCTCTCAACAAAAATGGAGGGGAG	2803
Qy	2613	ATG	CTAGAAATCCCCCTAGACGGAGGAAAAACCAATTTTATTCAGTGAATATACACATCTCTCT	2672
Db	2804	ATG	CTAGAAATCCCCCTAGACGGAGGAAAAACCAATTTTATTCAGTGAATATACACATCTCTCT	2863
Qy	2673	TGTT	CTTAAAAAGCAAGTGTCTTTTGGTGTGGAGGACAAAAATCCCTTACCATTTTCCAC	2732
Db	2864	TGTT	CTTAAAAAGCAAGTGTCTTTTGGTGTGGAGGACAAAAATCCCTTACCATTTT-CAC	2922
Qy	2733	GTGT	GCTACTAGAGATCTCAAAATATATAGTCTTGTTCGGGACCCCTTCATAGTACACCT	2792
Db	2923	GTGT	GCTACTAGAGATCTCAAAATATATAGTCTTGTTCGGGACCCCTTCATAGTACACCT	2982
Qy	2793	TAG	CGCTGAGACTGAGCCAGCTTGGGGGTGAGGTAGGTAGAGCCCTGTGTTAGGACAGAGCC	2852

## RESULT 5

US-10-377-072-25  
; Sequence 25, Application US/10377072  
; Publication NO. US20040009501A1  
; GENERAL INFORMATION:  
; APPLICANT: Millennium Pharmaceuticals Inc.  
; APPLICANT: Curtis, Rory A.J.  
; APPLICANT: Logan, Thomas Joseph  
; APPLICANT: Glucksmann, Maria A.  
; APPLICANT: Meyers, Rachel E.  
; APPLICANT: Williamson, Mark J.  
; APPLICANT: Rudolph-Owen, Laura A.  
; APPLICANT: Chun, Miyoung  
; APPLICANT: Tsai, Fong-Ying  
; TITLE OF INVENTION: NOVEL 25869, 25934, 26335, 50365, 21117.  
; TITLE OF INVENTION: 38652, 46508, 16816, 18839, 49937, 49931 AND 49933 MOLECULES  
; TITLE OF INVENTION: AND USSES THEREFOR  
; FILE REFERENCE: MPI03-018CMNM  
; CURRENT APPLICATION NUMBER: US/10/377,072  
; CURRENT FILING DATE: 2003-02-27  
; PRIOR APPLICATION NUMBER: US 09/895,860  
; PRIOR FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: US 60/215,370  
; PRIOR FILING DATE: 2000-06-29  
; PRIOR APPLICATION NUMBER: US 09/723,806  
; PRIOR FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: US 60/187,455  
; PRIOR FILING DATE: 2000-03-07  
; PRIOR APPLICATION NUMBER: US 09/843,297  
; PRIOR FILING DATE: 2001-04-25  
; PRIOR APPLICATION NUMBER: US 60/199,801  
; PRIOR FILING DATE: 2000-04-26  
; PRIOR APPLICATION NUMBER: US 09/861,801  
; PRIOR FILING DATE: 2001-05-21  
; PRIOR APPLICATION NUMBER: US 60/205,508  
; PRIOR FILING DATE: 2000-05-19  
; PRIOR APPLICATION NUMBER: US 09/816,494

PRIOR FILING DATE: 2001-03-23  
PRIOR APPLICATION NUMBER: US 09/815,419  
PRIOR FILING DATE: 2001-03-22  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 114  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 25  
LENGTH: 3544  
TYPE: DNA  
ORGANISM: Homo Sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (589)... (2586)  
US-10-377-072-25

Query Match 88.5%; Score 2950; DB 15; Length 3544;  
Best Local Similarity 95.0%; Pred. No. 0;  
Matches 3135; Conservative 0; Mismatches 0; Indels 165; Gaps 2;  
197 GCTTTCAAGTCCAGTGTAAAGCTGTTGGAGCGCGGAGCAAGGTAAAGAAATGATGTAATG 256  
224 GCTTTCAAGTCCAGTGTAAAGCTGTTGGAGCGCGGAGCAAGGTAAAGAAATGATGTAATG 283  
257 CGCTGGCTGCTCCAAAGCATCTTTGTTGGAATGTTATCCAGTCATCTCTTTATGA 316  
284 CGCTGGCTGCTCCAAAGCATCTTTGTTGGAATGTTATCCAGTCATCTCTTTATGA 343  
317 ATCAAAATGTAGGGGCTGCTTTGTGGAGCGAGTCTTTGCAAGAGCACATCAACGGGAAA 376  
344 ATCAAAATGTAGGGGCTGCTTTGTGGAGCGAGTCTTTGCAAGAGCACATCAACGGGAAA 403  
377 GAGAAAGAGACATTCACCTTGAGGGCTCTTGCTGAAATGTTGTTAACTCTCTCTTTGCC 436  
404 GAGAAAGAGACATTCACCTTGAGGGCTCTTGCTGAAATGTTGTTAACTCTCTCTTTGCC 463  
437 ACTCAACCAAGCTGACCTCATACACTTTAGTACAAATGAGTGGCTGAGCTTTGAGC 496  
464 AGTCAACCAAGCTGACCTCATACACTTTAGTACAAATGAGTGGCTGAGCTTTGAGC 523  
497 ACACCAACCATTAACATCATCGTGGCAATTAAGAGAGGAGTGGGAAAGAGGACTTTATTG 556  
524 ACACCAACCATTAACATCATCGTGGCAATTAAGAGAGGAGTGGGAAAGAGGACTTTATTG 583  
557 TTGTCATGCCCCATGAGATGATGGAATCAAAATGTTACTGAGAGTGTGGTGGCTCTGC 616  
584 TTGTCATGCCCCATGAGATGATGGAATCAAAATGTTACTGAGAGTGTGGTGGCTCTGC 643  
617 TGGAAAGTGGAAACGGAAGGCTGCTGCTAAATGATAGCGGCCAATTTGTGGAATACAATA 676  
644 TGGAAAGTGGAAACGGAAGGCTGCTGCTAAATGATAGCGGCCAATTTGTGGAATACAATA 703  
677 CATCCCAATTTTGAAGCCATTAATATCAACTGCTCCAAAGCTATGAAGCGAGGTTGC 736  
704 CATCCCAATTTTGAAGCCATTAATATCAACTGCTCCAAAGCTATGAAGCGAGGTTGC 763  
737 AACAGGACAAAGTGTAAATACAGAGCTCATCCAGCATTCAGCGAAACATAAGGTTGACA 796  
764 AACAGGACAAAGTGTAAATACAGAGCTCATCCAGCATTCAGCGAAACATAAGGTTGACA 823  
797 TTGAATGCAATCAGAGGTTGATTTACGATCAAAAGCTCCCAAGATGTTGCCCTCTCTCT 856  
824 TTGAATGCAATCAGAGGTTGATTTACGATCAAAAGCTCCCAAGATGTTGCCCTCTCTCT 883  
857 CTTCAAGCTGTTTCTCACTGCTACTCTGTTGGAATGTTGGAAGAGCTTCAACTCTGTTTC 916  
884 CTTCAAGCTGTTTCTCACTGCTACTCTGTTGGAATGTTGGAAGAGCTTCAACTCTGTTTC 943  
917 ACCTGCTTGC----- 926  
944 ACCTGCTTGCAGGTGGGTTTGTCTGAGTTCTCTGTTTCCCTGGCCTCTGTGAAGGAA 1003  
927 ----- 926

Db 1004 AATCCACTCTAGTCCCTACCTACCTGCAATTTCTCAGCGCTTGCTTACCTGTGTGCCAATTTGGGC 1063  
QY 927 -----AGAGC 932  
Db 1064 CAACCCGAATTTCTCCCAATCTTTTATCTTGGCTGCCAGCGAGATGCTCTCAACAAGGAGC 1123  
QY 933 TGATGCAGCAGAAATGGATTTGTTATGTTTAAATGCCAGCAATACCTGTCCAAAGCCTG 992  
Db 1124 TGATGCAGCAGAAATGGATTTGTTATGTTTAAATGCCAGCAATACCTGTCCAAAGCCTG 1183  
QY 993 ACTTTATCCCGAGTCTCAATTTCTGCTGTGCTGTGAATGACAGCTTTTGTGAGAAA 1052  
Db 1184 ACTTTATCCCGAGTCTCAATTTCTGCTGTGCTGTGAATGACAGCTTTTGTGAGAAA 1243  
QY 1053 TTTTGGCGTGTGAGCAAAATCAGTGTATTTCAATGAGAAAGCAAAAGCCTCCAAATGGAT 1112  
Db 1244 TTTTGGCGTGTGAGCAAAATCAGTGTATTTCAATGAGAAAGCAAAAGCCTCCAAATGGAT 1303  
QY 1113 GHTTCTAGTCACTGTTTATAGTGGGATCTCCGCTCCGCCACCATCGTATCGCCTACA 1172  
Db 1304 GTGTTCTAGTCACTGTTTATAGTGGGATCTCCGCTCCGCCACCATCGTATCGCCTACA 1363  
QY 1173 TCATGAAGAGATGGAATGCTTTAGATGAAGTTACAGATTTGTGAAGAAAAGAC 1232  
Db 1364 TCATGAAGAGATGGAATGCTTTAGATGAAGTTACAGATTTGTGAAGAAAAGAC 1423  
QY 1233 CTACTATATCTCCAAATCTTCAATTTCTGGGCCAACTCTCTGGACTATGAGAAAGATTA 1292  
Db 1424 CTACTATATCTCCAAATCTTCAATTTCTGGGCCAACTCTCTGGACTATGAGAAAGATTA 1483  
QY 1293 AGAACAGACTGGAGCATCGGGCCAAAGAGCAAACTCAAGCTGTGACCTGGAGAGC 1352  
Db 1484 AGAACAGACTGGAGCATCGGGCCAAAGAGCAAACTCAAGCTGTGACCTGGAGAGC 1543  
QY 1353 CAAATGAACCTGCTCCCTGCTGCTCAGAGGCTGAGCAAGAAAGCAGAGCCCTCAGTC 1412  
Db 1544 CAAATGAACCTGCTCCCTGCTGCTCAGAGGCTGAGCAAGAAAGCAGAGCCCTCAGTC 1603  
QY 1413 CACCTGTGCGGACTCTGCTACCTCAGAGGACAGGCTGGAAGACAGCAATAAGCTCAAGCGTT 1472  
Db 1604 CACCTGTGCGGACTCTGCTACCTCAGAGGACAGGCTGGAAGACAGCAATAAGCTCAAGCGTT 1663  
QY 1473 GCGTGGCCAGCGTGGCCAGCGTGCAGCGCTGCTGTAGAGGACAGCGCTGGTACAGG 1532  
Db 1664 GCGTGGCCAGCGTGGCCAGCGTGCAGCGCTGCTGTAGAGGACAGCGCTGGTACAGG 1723  
QY 1533 CGCTCAGTGGGCTGCACTGTCCGCGAGACAGGCTGGAAGACAGCAATAAGCTCAAGCGTT 1592  
Db 1724 CGCTCAGTGGGCTGCACTGTCCGCGAGACAGGCTGGAAGACAGCAATAAGCTCAAGCGTT 1783  
QY 1593 CTTTCTCTGGAATATCAAAATCAGTTTATATTCAGCGAGCATGGAGCATCTCTTACATG 1652  
Db 1784 CTTTCTCTGGAATATCAAAATCAGTTTATATTCAGCGAGCATGGAGCATCTCTTACATG 1843  
QY 1653 GCTTCTCTCATCAGAAGATGCTTTGGAATACCTCAAACTCTTCCACTACTCTGATGGGA 1712  
Db 1844 GCTTCTCTCATCAGAAGATGCTTTGGAATACCTCAAACTCTTCCACTACTCTGATGGGA 1903  
QY 1713 CCAACAAGCTATGCAAGTCTCCCTGTTTCAAGAACTATCGGAGCAGATTCGCCGAAACCA 1772  
Db 1904 CCAACAAGCTATGCAAGTCTCCCTGTTTCAAGAACTATCGGAGCAGATTCGCCGAAACCA 1963  
QY 1773 GTCTGTATAGGAGGAGCCAGCATCCCAAGAGCTGAGACCGCCAGCGCTTCAGACA 1832  
Db 1964 GTCTGTATAGGAGGAGCCAGCATCCCAAGAGCTGAGACCGCCAGCGCTTCAGACA 2023  
QY 1833 GCCAGAGCAAGCGATTTGCAATTCGGTTCAGAACTCCAGAGCAGTGGCACCGCCAGAGGTCCC 1892  
Db 2024 GCCAGAGCAAGCGATTTGCAATTCGGTTCAGAACTCCAGAGCAGTGGCACCGCCAGAGGTCCC 2083  
QY 1893 TTTTATCTCACTGATCGAAGTGGGAGCGTGGAGGACAAATACACACAGCTTCTCTTT 1952  
Db 2084 TTTTATCTCACTGATCGAAGTGGGAGCGTGGAGGACAAATACACACAGCTTCTCTTT 2143



334 CTATCGGAGCAGACTCCCGAACAACAGTCTCTGATAGGAGGAGCCAGCATCCCAAGAAG 393  
1808 CTGACAGACCGCAGCGCTTCAGACAGCCAGAGCGATTCGATTCGGTTCAGAACCCAGC 1867  
394 CTGACAGACCGCAGCGCTTCAGACAGCCAGAGCGATTCGATTCGGTTCAGAACCCAGC 453  
1868 AGCAGTGGCAGCCGCGCCAGAGTCCCTTTTATCTCCACTGCTGATCGAAGTGGAGCGTGGAG 1927  
454 AGCAGTGGCAGCCGCGCCAGAGTCCCTTTTATCTCCACTGCTGATCGAAGTGGAGCGTGGAG 513  
1928 GACAAATTACACACAGCTTCTCTTTTGGCCCTTTCCACACGCGCAGCAGCACCTCAGGAAG 1987  
514 GACAAATTACACACAGCTTCTCTTTTGGCCCTTTCCATCAGCCAGCAGCACCTCAGGAAG 573  
1988 TGTGCTGGCTGGCGCTTAAGGGCTGGCAGCTCGATATCTTGGCCCGCCAGACCTCTACC 2047  
574 TGTGCTGGCGCTGGCGCTTAAGGGCTGGCAGCTCGATATCTTGGCCCGCCAGACCTCTACC 633  
2048 CTTTCCCTGATCAGCAGCTGGTATTTTGGCCACAGAGTCCCTCACACTTCTACTCTGCTCA 2107  
634 CTTTCCCTGATCAGCAGCTGGTATTTTGGCCACAGAGTCCCTCACACTTCTACTCTGCTCA 693  
2108 GCCATCTAGGAGGAGTGCAGATTACTCTGCTAGAGTGCAGCAGCTGCCCCACTTGC 2167  
694 GCCATCTAGGAGGAGTGCAGATTACTCTGCTAGAGTGCAGCAGCTGCCCCACTTGC 753  
2168 GGAGACCAAGTCTATTCTGTGCGAGCGCGCAGAGCCAAAGTGCAGAGCTGACTCGCGG 2227  
754 GGAGACCAAGTCTATTCTGTGCGAGCGCGCAGAGCCAAAGTGCAGAGCTGACTCGCGG 813  
2228 CGAGCTGGCTAGAGAGAGCGCCCTTTGAAAGCAGTTTAAAGCAGAGCTGCCAAATG 2287  
814 CGAGCTGGCTAGAGAGAGCGCCCTTTGAAAGCAGTTTAAAGCAGAGCTGCCAAATG 873  
2288 GAATTTGAGAGAGCATCATCTCAGAGAACAGTGCAGGAGAGCTGGGAAAGTGGGC 2347  
874 GAATTTGAGAGAGCATCATCTCAGAGAACAGTGCAGGAGAGCTGGGAAAGTGGGC 933  
2348 AGTCAGTCTAGCTTTTGGGCGAGCATGGAATCAATTGAGTCTCCTGAGAGAAAGACAC 2407  
934 AGTCAGTCTAGCTTTTGGGCGAGCATGGAATCAATTGAGTCTCCTGAGAGAAAGACAC 993  
2408 TTGTCAGCTCTATAGACAAATTTTCTTCTGTCACAAAATAATCCCTGTAATCTGA 2467  
994 TTGTCAGCTCTATAGACAAATTTTCTTCTGTCACAAAATAATCCCTGTAATCTGA 1053  
2468 AATATATATGTACATACATATATTTTGGAAATGAGCTATGCTGTAAGAACAC 2527  
1054 AATATATATGTACATACATATATTTTGGAAATGAGCTATGCTGTAAGAACAC 1113  
2528 AGTGGATCAACCCAGTTGTTACTCTCTTAACATCTGATCTGAGATCAGCTAATCT 2587  
1114 AGTGGATCAACCCAGTTGTTACTCTCTTAACATCTGATCTGAGATCAGCTAATCT 1173  
2588 TCTCTCAACAAAATGGAAGCGCAGATGCTAGATCCCTCAGCGAGGAAACCAT 2647  
1174 TCTCTCAACAAAATGGAAGCGCAGATGCTAGATCCCTCAGCGAGGAAACCAT 1233  
2648 TTATTCAGTGAATTACATCCTCTGTTCTTAAAAAGCAAGTGTCTTTGGTGTGGAG 2707  
1234 TTATTCAGTGAATTACATCCTCTGTTCTTAAAAAGCAAGTGTCTTTGGTGTGGAG 1293  
2708 GACAAAATCCCTACCATTTTCCAGGTTGCTACTAAGAGTCTCAATATTAGTCTTT 2767  
1294 GACAAAATCCCTACCATTTTCCAGGTTGCTACTAAGAGTCTCAATATTAGTCTTT 1352  
2768 GTCCGAGCCCTTCCATAGTACACCTTAGCGCTGAGCTGAGCCAGCTTGGGGTCAAGTA 2827  
1353 GTCCGAGCCCTTCCATAGTACACCTTAGCGCTGAGCTGAGCCAGCTTGGGGTCAAGTA 1412  
2828 GGTAGACCCCTGTTAGGAGCAGAGCCTTAGTGAATCCAAAGAGAAATGATCTTATCCAA 2887  
1413 GGTAGACCCCTGTTAGGAGCAGAGCCTTAGTGAATCCAAAGAGAAATGATCTTATCCAA 1472

2888 GCTGATTCAAAAACCCACGCTCACCTGACAGCCGAGGGAACAGAGCATCACTCTGCTGGA 2947  
1473 GCTGATTCAAAAACCCACGCTCACCTGACAGCCGAGGGAACAGAGCATCACTCTGCTGGA 1532  
2948 CGGACCATTAAGGGGCTTGGCAAGGCTCTACCTTAGAGCAAAACCCAGTACCTCAGACAGGA 3007  
1533 CGGACCATTAAGGGGCTTGGCAAGGCTCTACCTTAGAGCAAAACCCAGTACCTCAGACAGGA 1592  
3008 AAGTCGGGGCTTTGACCACTACCAATCTGTAAGCCCAATTTCTAGGCATTTGTAATAGG 3067  
1593 AAGTCGGGGCTTTGACCACTACCAATCTGTAAGCCCAATTTCTAGGCATTTGTAATAGG 1652  
3068 TAGTCTAGCTAGTCACTTTTTCAGACCAATTTCAAACTGTCTATGCACAAAATTTCCCGTGG 3127  
1653 TAGTCTAGCTAGTCACTTTTTCAGACCAATTTCAAACTGTCTATGCACAAAATTTCCCGTGG 1712  
3128 GCCTAGATGGAGATAAATTTTCTTCTCAGCTTTATGAAGAGAGGAAACTCTCTTA 3187  
1713 GCCTAGATGGAGATAA-TTTTTTTTCTTCTCAGCTTTATGAAGAGAGGAAACTCTCTTA 1771  
3188 GGATTCAGCTGAACCCAGCAACCTGGCAACATCAGGATTTAAGCTAAGGTTGGGAGGC 3247  
1772 GGATTCAGCTGAACCCAGCAACCTGGCAACATCAGGATTTAAGCTAAGGTTGGGAGGC 1831  
3248 TAAGCAGTCTACCTCCCTCTTTGTAATCAAGAAATTTTAAAAATGGGATTTGCAATCC 3307  
1832 TAAGCAGTCTACCTCCCTCTTTGTAATCAAGAAATTTTAAAAATGGGATTTGCAATCC 3307  
3308 TTTAAATAAAGATGAATCTGGTTTC 3332  
1892 TTTAAATAAAGATGAATCTGGTTTC 1916

## RESULT 7

US-10-094-749-673  
; Sequence 673, Application US/10094749  
; Publication No. US20030219741A1  
; GENERAL INFORMATION:  
; APPLICANT: ISOGAI, TAKAO  
; APPLICANT: SUGIYAMA, TOMOYASU  
; APPLICANT: OTSUKI, TETSUJI  
; APPLICANT: WAKAMATSU, AI  
; APPLICANT: SATO, HIROYUKI  
; APPLICANT: ISHII, SHIZUKO  
; APPLICANT: YAMAMOTO, JUN-ICHI  
; APPLICANT: ISONO, YUUKO  
; APPLICANT: HIO, YURI  
; APPLICANT: OTSUKA, KAORU  
; APPLICANT: NAGAI, KEIICHI  
; APPLICANT: IRIE, RYOTARO  
; APPLICANT: TAMECHIKA, ICHIRO  
; APPLICANT: SEKI, NAOHICO  
; APPLICANT: YOSHIKAWA, TSUTOMU  
; APPLICANT: OTSUKA, MOTOKYUKI  
; APPLICANT: NAGAHARI, KENJI  
; APPLICANT: MASUHO, YASUHIKO  
; TITLE OF INVENTION: NOVEL FULL-LENGTH cDNA  
; FILE REFERENCE: 084335/0160  
; CURRENT APPLICATION NUMBER: US/10/094,749  
; PRIOR FILING DATE: 2002-03-12  
; PRIOR APPLICATION NUMBER: 60/350,435  
; PRIOR FILING DATE: 2002-01-24  
; PRIOR APPLICATION NUMBER: JP 2001-328381  
; NUMBER OF SEQ ID NOS: 3381  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 673  
; LENGTH: 2102  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-094-749-673

Query Match	52.9%	Score 1762.4	DB 15	Length 2102
Best Local Similarity	92.2%	Mismatches 0	Indels 154	Gaps 1
Matches 1937	Conservative	0	Mismatches	
QY	507	TACATCATCGTGGCAAATTAAAGAGAGAGTGGAAAGAGGACACTTATGTTGTGTCATGGC	566	
DB	1	TACATCATCGTGGCAAATTAAAGAGAGGTTGGAAAGAGGACTTATGTTGTGTCATGGC	60	
QY	567	CCATGAGATGATTGGAACTCAAAATTGTTACTGAGAGGTTGTTGGCTCTGCTGGAAAGTGG	626	
DB	61	CCATGAGATGATTGGAACTCAAAATTGTTACTGAGAGGTTGTTGGCTCTGCTGGAAAGTGG	120	
QY	627	AACGGAAAAGTGTGCTAAATTCATAGCGCGGCATTGTTGGAATACAAATACATCCACAT	686	
DB	121	AACGGAAAAGTGTGCTAAATTCATAGCGCGGCATTGTTGGAATACAAATACATCCACAT	180	
QY	687	TTTGAAGCCATTAAATATCAACTGCTCCAAAGCTTATGAAGCGAAGTTGCAACAGGCAA	746	
DB	181	TTTGAAGCCATTAAATATCAACTGCTCCAAAGCTTATGAAGCGAAGTTGCAACAGGCAA	240	
QY	747	AGGTAAATTACAGAGCTCATCAGCATTTACCGGAAACATAAGGTTGACATTGATTCGAG	806	
DB	241	AGGTAAATTACAGAGCTCATCAGCATTTACCGGAAACATAAGGTTGACATTGATTCGAG	300	
QY	807	TCAGAAGGTTGTAGTTTACGATCAAAAGCTCCCAAGATGTGCCTCTCTCTCTTCCAGACTG	866	
DB	301	TCAGAAGGTTGTAGTTTACGATCAAAAGCTCCCAAGATGTGCCTCTCTCTCTTCCAGACTG	360	
QY	867	TTTTCTCACTGTACTTCTGGTAAACTGAGAAAGAGCTTCAACTCTGTTCACTGCTGTTGC	926	
DB	361	TTTTCTCACTGTACTTCTGGTAAACTGAGAAAGAGCTTCAACTCTGTTCACTGCTGTTGC	420	
QY	927	-----	926	
DB	421	AGTGGGTTTGTGAGTTCTCTCGTTGTTCCCTGSCCTCTGTGAAGAAAATCCACTCT	480	
QY	927	-----	926	
DB	481	AGTCCCTACCTGCATTCTTCAGCGCTTGCTTACTGTTGCCAACATTGGGCCCAACCTTAAT	540	
QY	927	-----	926	
DB	541	TCTTCCCAATCTTTATCTTGGCTGCCAGAGATGTCCTCAACAGAGGAGCTGATCGAGCA	600	
QY	943	GAATGGGATTGGTTATGTGTTAAATGCCAGCAATACCTGTCCAAAGCCTGACTTTATCCC	1002	
DB	601	GAATGGGATTGGTTATGTGTTAAATGCCAGCAATACCTGTCCAAAGCCTGACTTTATCCC	660	
QY	1003	CGAGTCTCATTTCTGCTGGTGTGCTGTGAATGACAGCTTTTGTGAGAAAATTTTGGCGTG	1062	
DB	661	CGAGTCTCATTTCTGCTGGTGTGCTGTGAATGACAGCTTTTGTGAGAAAATTTTGGCGTG	720	
QY	1063	GTTGACCAAAATCAGTAGATTTCATTGAGAAAAGCAAAGCCTCCAATGGATGTGTTCTAGT	1122	
DB	721	GTTGACCAAAATCAGTAGATTTCATTGAGAAAAGCAAAGCCTCCAATGGATGTGTTCTAGT	780	
QY	1123	GCACTGTTTACCTGGATCTCCGGCTCCGGCCACCATCGCTATCGGCTACATCATGTAAGAG	1182	
DB	781	GCACTGTTTACCTGGATCTCCGGCTCCGGCCACCATCGCTATCGGCTACATCATGTAAGAG	840	
QY	1183	GATGACATGTCTTTAGATGAAGCTTACAGATTTGTGAAGAAAAGAACCCCTACTATATC	1242	
DB	841	GATGACATGTCTTTAGATGAAGCTTACAGATTTGTGAAGAAAAGAACCCCTACTATATC	900	
QY	1243	TCCAAACTTCAATTTTCTGGGCCAACTCTTGGACTATGAGAAAGAGATTAAAGAACAGAC	1302	
DB	901	TCCAAACTTCAATTTTCTGGGCCAACTCTTGGACTATGAGAAAGAGATTAAAGAACAGAC	960	
QY	1303	TGGAGCATCGGGCCAAAGAGCAAACTCAAGCTGTGACCTGGAGAGCCAAATGAACC	1362	
DB	961	TGGAGCATCGGGCCAAAGAGCAAACTCAAGCTGTGACCTGGAGAGCCAAATGAACC	1020	
QY	1363	TGTCCTCTGCTCTCAGAGGGTGGACAGAAAAGCGAGACGCCCTCTCAGCTCCACCTGTGC	1422	

1021	DB	TGTCCCTGCTGTCTCAGAGGGTGGACAGAAAAGCAGACGCCCTCAGTCCACCCTGTGC	1080
1423	QY	CGACTCTGCTACCTCAGAGCAGCAGGACAAAGGCCCGTGTGATCCCGCAGCGTCCCGAG	1482
1081	DB	CGACTCTGCTACCTCAGAGCAGCAGGACAAAGGCCCGTGTGATCCCGCAGCGTCCCGAG	1140
1483	QY	CGTGCACAGCGTGAGCCGTGCTGTGTAGAGACAGCCGCGTGTGTACAGGCGCTCAGTGG	1542
1141	DB	CGTGCACAGCGTGAGCCGTGCTGTGTAGAGACAGCCGCGTGTGTACAGGCGCTCAGTGG	1200
1543	QY	GCTGCACCTGTCCGACAGCAGGCTGGAACAGCAGCAATAAGCTCAAGCGTTCCTTCTCTCT	1602
1201	DB	GCTGCACCTGTCCGACAGCAGGCTGGAACAGCAGCAATAAGCTCAAGCGTTCCTTCTCTCT	1260
1603	QY	GGATATCAAAATCAGTTTTCATATTACGCCAGCATGGCAGGATTCCTTACATGCGTTCCTCTC	1662
1261	DB	GGATATCAAAATCAGTTTTCATATTACGCCAGCATGGCAGGATTCCTTACATGCGTTCCTCTC	1320
1663	QY	ATCAGAGATGCTTTTGGAAATPACTACAAACCTTTCCACTACTCTGATGGAGCAACAAAGCT	1722
1321	DB	ATCAGAGATGCTTTTGGAAATPACTACAAACCTTTCCACTACTCTGATGGAGCAACAAAGCT	1380
1723	QY	ATGCCAGTTCTCCCTGTTTCAGGAATATCGAGCAGACTCCCGAAACAGTCTCTGTATAA	1782
1381	DB	ATGCCAGTTCTCCCTGTTTCAGGAATATCGAGCAGACTCCCGAAACAGTCTCTGTATAA	1840
1783	QY	GGAGGAGCCAGCATCCCCAAGAAGCTGCAGACCCGAGGCGCTTCAGACAGCCAGAGCAA	1842
1441	DB	GGAGGAGCCAGCATCCCCAAGAAGCTGCAGACTGCCAGGCGCTTCAGACAGCCAGAGCAA	1500
1843	QY	CGCATATGCATTCCGTCAGAACACAGCAGCAGTGGCACCGCCAGAGGTCCTTTTATCTCC	1902
1501	DB	CGCATATGCATTCCGTCAGAACACAGCAGCAGTGGCACCGCCAGAGGTCCTTTTATCTCC	1560
1903	QY	ACTGCATCGAAGTGGAGCGTGGAGACAAATTACACACCCAGCTTCCTTTTCGCGCTTTC	1962
1561	DB	ACTGCATCGAAGTGGAGCGTGGAGACAAATTACACACCCAGCTTCCTTTTCGCGCTTTC	1620
1963	QY	CACGAGCCAGCAGCACTCACGAAGTCTGCTGGCGCTGGGCGCTTAAGGGCTGGCACTCGGA	2022
1621	DB	CACGAGCCAGCAGCACTCACGAAGTCTGCTGGCGCTGGGCGCTTAAGGGCTGGCACTCGGA	1680
2023	QY	TATCTTGGCCCCCAGACCTCTACCCCTTCCCTGTACACAGCAGCTGGTATTTTGGCCACAGA	2082
1681	DB	TATCTTGGCCCCCAGACCTCTACCCCTTCCCTGTACACAGCAGCTGGTATTTTGGCCACAGA	1740
2083	QY	GTCTCTCACATTTCTACTCTGCTCAGCCATCTACGGAGCAGTGCAGTAACTCTGCGCTA	2142
1741	DB	GTCTCTCACATTTCTACTCTGCTCAGCCATCTACGGAGCAGTGCAGTAACTCTGCGCTA	1800
2143	QY	CAGCTGCAGCAGCTGCCCACTTGGGAGACCAAGTCTATTCTGTGCGCAGCGCGCAGAA	2202
1801	DB	CAGCTGCAGCAGCTGCCCACTTGGGAGACCAAGTCTATTCTGTGCGCAGCGCGCAGAA	1860
2203	QY	GCCAAGTGACAGACTGACTCGCGGCGGAGCTGGCATGAAGAGACCCCTTTGAAAGCA	2262
1861	DB	GCCAAGTGACAGACTGACTCGCGGCGGAGCTGGCATGAAGAGACCCCTTTGAAAGCA	1920
2263	QY	GTTTTAAACGACAGAGCTGCCAAATGGAAATTTGGAGAGAGCATCATGTACAGAAACAGGTC	2322
1921	DB	GTTTTAAACGACAGAGCTGCCAAATGGAAATTTGGAGAGAGCATCATGTACAGAAACAGGTC	1980
2323	QY	ACGGAGAGAGCTGGGAAAGTGGGAGTCAGTCTAGCTTTTCGGGCGAGCATGGAAATCAT	2382
1981	DB	ACGGAGAGAGCTGGGAAAGTGGGAGTCAGTCTAGCTTTTCGGGCGAGCATGGAAATCAT	2040
2383	QY	TGAGGTCTCTTGAGAGAAAGACACTTGTGACTTCTATAGACAAATTTTTTTTCTGTGTC	2442
2041	DB	TGAGGTCTCTTGAGAGAAAGACACTTGTGACTTCTATAGACAAATTTTTTTTCTGTGTC	2100
2443	QY	AC 2444	



Db 2101 AC 2102

RESULT 8

US-10-072-012-255  
; Sequence 255, Application US/10072012  
; Publication No. US2004003493A1

GENERAL INFORMATION:

; APPLICANT: Tchernev, Velizar  
; APPLICANT: Spytek, Kimberly  
; APPLICANT: Zerhusen, Bryan  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Shimkets, Richard  
; APPLICANT: Li, Li  
; APPLICANT: Gangolli, Beha  
; APPLICANT: Padigaru, Muralidhara  
; APPLICANT: Anderson, David W.  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Miller, Charles E.  
; APPLICANT: Gerlach, Valerie  
; APPLICANT: Taupier, Jr, Raymond J.  
; APPLICANT: Gusev, Vladimir Y.  
; APPLICANT: Colman, Steven D.  
; APPLICANT: Wolenc, Adam R.  
; APPLICANT: Pena, Carol E. A  
; APPLICANT: Furtak, Katarzyna  
; APPLICANT: Grosse, William M.  
; APPLICANT: Alsobrook II, John P.  
; APPLICANT: Lepley, Denise M.  
; APPLICANT: Riegez, Daniel K.  
; APPLICANT: Burgess, Catharine E.

; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same

; FILE REFERENCE: 21402-258

; CURRENT APPLICATION NUMBER: US/10/072,012

; CURRENT FILING DATE: 2002-01-31

; PRIOR APPLICATION NUMBER: 60/265,102

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: 60/265,514

; PRIOR FILING DATE: 2001-01-31

; PRIOR APPLICATION NUMBER: 60/265,517

; PRIOR FILING DATE: 2001-01-31

; PRIOR APPLICATION NUMBER: 60/265,412

; PRIOR FILING DATE: 2001-01-31

; PRIOR APPLICATION NUMBER: 60/265,395

; PRIOR FILING DATE: 2001-01-31

; PRIOR APPLICATION NUMBER: 60/266,406

; PRIOR FILING DATE: 2001-02-02

; PRIOR APPLICATION NUMBER: 60/266,767

; PRIOR FILING DATE: 2001-02-05

; PRIOR APPLICATION NUMBER: 60/267,057

; PRIOR FILING DATE: 2001-02-07

; PRIOR APPLICATION NUMBER: 60/266,975

; PRIOR FILING DATE: 2001-02-07

; PRIOR APPLICATION NUMBER: 60/267,459

; PRIOR FILING DATE: 2001-02-08

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 1391

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 255

; LENGTH: 2200

; TYPE: DNA

; ORGANISM: Homo sapiens

JS-10-072-012-255

Query Match 52.3%; Score 1742; DB 12; Length 2200;

Best Local Similarity 90.5%; Pred. No. 0;

Matches 1991; Conservative 0; Mismatches 0; Indels 209; Gaps 4;

502 ACCATTACATCATCGTGGGCAAAATTAAGAGAGAGGTGGGAAAGAGAGCTTATTGTTGTC 561

1 ACCATTACATCATCGTGGGCAAAATTAAGAGAGAGGTGGGAAAGAGAGCTTATTGTTGTC 60

562 ATGCCCATGAGATGATTGGAACCTCAATTTGTTACTGAGAGGTGGTGGCTCTGCTGAA 621

Db 61 ATGCCCATGAGATGATTGGAACCTCAATTTGTTACTGAGAGGTGGTGGCTCTGCTGAA 120

Qy 622 AGTGGAAACGAAAAAGTGTCTTAATTGATAGCGCCATTTGTGTAATACATAATACATCC 681

Db 121 AGTGGAAACGAAAAAGTGTCTTAATTGATAGCGCCATTTGTGTAATACATAATACATCC 180

Qy 682 CACATTTTGGAAAGCCATTAAATCAACTGCTCCAAAGCTTATGAAGCGAAGTTGCAACAG 741

Db 181 CACATTTTGGAAAGCCATTAAATCAACTGCTCCAAAGCTTATGAAGCGAAGTTGCAACAG 240

Qy 742 GACAAAGTGTAAATTACAGAGCTCATCCAGCAATTCAGCGAACATA-----AG 789

Db 241 GACAAAGTGTAAATTACAGAGCTCATCCAGCAATTCAGCGAACATAAGGTAACCGCTCAG 300

Qy 790 GTTGACATTTGATTCAGAGTCAGAAAGTTGTAGTTTACGATCAAAAGCTCCCAAGATTTGCC 849

Db 301 GTTGACATTTGATTCAGAGTCAGAAAGTTGTAGTTTACGATCAAAAGCTCCCAAGATTTGCC 360

Qy 850 TCTCTCTCTTCAGACTGTTTCTCACTGTAATCTTCTGGGTAAACTGGAGAGAGCTTCAAC 909

Db 361 TCTCTCTCTTCAGACTGTTTCTCACTGTAATCTTCTGGGTAAACTGGAGAGAGCTTCAAC 420

Qy 910 TCTGTTCACTGCTTGC----- 926

Db 421 TCTGTTCACTGCTTGCAGGTTTATCTTAGTGGGTTTGTGAGTCTCTCGTGTGTTTC 480

Qy 927 ----- 926

Db 481 CTTGGCTCTGTGAAGGAAATCCACTAGTCCCTACCTGCAATTTCTCAGCCTTGCTTA 540

Qy 927 ----- 926

Db 541 CTTGTTGCCAATTTGGGCCAACCCGAATCTTCCCAATCTTATCTTGGCTGCCAGCA 600

Qy 927 -----AGAGCTGATGCAGAGAAATGGATGGTGTATGTGTTAAATGCCAGC 973

Db 601 GATGTCTCTCAACAGAGAGCTGATGCAGAGAAATGGATGGTGTATGTGTTAAATGCCAGC 660

Qy 974 ATACCTGTCCAAAGCCGTGACTTTATCCCGAGTCTCATTTCCCTGCGTGCCTGTGTAAT 1033

Db 661 ATACCTGTCCAAAGCCGTGACTTTATCCCGAGTCTCATTTCCCTGCGTGCCTGTGTAAT 720

Qy 1034 GACAGCTTTTGTGAGAAAAATTTTGGCGTGGTGGACAAATCAGTAGATTTTCATT----- 1087

Db 721 GACAGCTTTTGTGAGAAAAATTTTGGCGTGGTGGACAAATCAGTAGATTTTCATTGGTAAG 780

Qy 1088 -----GAGAAAGCAAAAGCCCTCAATGGATGTGTCTTAGTGCATGTTTAGCT 1135

Db 781 TTGACTTATACAGAGAAAGCAAAAGCCCTCAATGGATGTGTCTTAGTGCATGTTTAGCT 840

Qy 1136 GGGATCTCCGCTCCGCCACCATCGCTATCGCTACATCATGAAGAGGATGGATGCT 1195

Db 841 GGGATCTCCGCTCCGCCACCATCGCTATCGCTACATCATGAAGAGGATGGATGCT 900

Qy 1196 TTAGATGAAGCTTAC---AGATTTGTGAAAGAAAAAGACCTACTATATCTCCAAATTC 1252

Db 901 TTAGATGAAGCTTACAGGAGATTTGTGAAAGAAAAAGACCTACTATATCTCCAAATTC 960

Qy 1253 AATTTTCTGGGCCAACTCTCGACTATGAGAGAGATTTAAGAACCCAGACTGGAGCATCA 1312

Db 961 AATTTTCTGGGCCAACTCTCGACTATGAGAGAGATTTAAGAACCCAGACTGGAGCATCA 1020

Qy 1313 GGGCCAAAGAGCAAACTCAAGCTGCTGCACCTGGAGAAAGCAAAATGAACCTGTCCCTGCT 1372

Db 1021 GGGCCAAAGAGCAAACTCAAGCTGCTGCACCTGGAGAAAGCAAAATGAACCTGTCCCTGCT 1080

Qy 1373 GTCTCAGAGGTGGACAGAAAGCCAGAGCGCCCTCAGTCCACCTGTGCGGACTGCT 1432

Db 1081 GTCTCAGAGGTGGACAGAAAGCCAGAGCGCCCTCAGTCCACCTGTGCGGACTGCT 1140

Qy 1433 ACCTCAGAGCGAGCAGGACAAAGCGCCGTGATCCGCGAGCGTGCAGCGTCCCGAGCGTCCCGCAGC 1492

Db 1141 ACCTCAGAGCAGGACAAAGGCCGCTGCATCCGCCAGCGTGCCAGCGTCCCGCAGC 1200  
Qy 1493 GTGACCCGCTCGCTGTAGAGGACAGCCCGCTGTGTACAGGCGCTCAGTGGGCTGCACCTG 1552  
Db 1201 GTGACCCGCTCGCTGTAGAGGACAGCCCGCTGTGTACAGGCGCTCAGTGGGCTGCACCTG 1260  
Qy 1553 TCCGACAGCAGCTGGAGACAGCAATAGCTCAGGCTTCTCTCTCTGGATATCAAA 1612  
Db 1261 TCCGACAGCAGCTGGAGACAGCAATAGCTCAGGCTTCTCTCTCTGGATATCAAA 1320  
Qy 1613 TCAGTTTCATATTCAGCCAGCAGTGCAGCATCTTACATGCTCTCTCTCATCAGAGAT 1672  
Db 1321 TCAGTTTCATATTCAGCCAGCAGTGCAGCATCTTACATGCTCTCTCTCATCAGAGAT 1380  
Qy 1673 GCTTTGGATATCAAACTTCCTACTCTCTGGATGGAGCAACAGCTATGCAGTTC 1732  
Db 1381 GCTTTGGATATCAAACTTCCTACTCTCTGGATGGAGCAACAGCTATGCAGTTC 1440  
Qy 1733 TCCCTGTTCAGGAAGTATCGGAGCAGTCCCGCAAGCAGTCTCTGATAAGGAGAGGCC 1792  
Db 1441 TCCCTGTTCAGGAAGTATCGGAGCAGTCCCGCAAGCAGTCTCTGATAAGGAGAGGCC 1500  
Qy 1793 AGCATCCCAAGAGCTGCAGACCGCAGGCTTTCAGACAGCAGAGCAAGCGATTGCAT 1852  
Db 1501 AGCATCCCAAGAGCTGCAGACCGCAGGCTTTCAGACAGCAGAGCAAGCGATTGCAT 1560  
Qy 1853 TCGGTCAGAACAGCAGCAGTGCAGCCGCGCAGAGCTCTCTTTTATCTCCACTCATCGA 1912  
Db 1561 TCGGTCAGAACAGCAGCAGTGCAGCCGCGCAGAGTCTCTTTTATCTCCACTCATCGA 1820  
Qy 1913 AGTGGAGCGTGGAGGACAAATACACACAGCTCTCTTTTGGCTTTTCCACAGCGAG 1972  
Db 1621 AGTGGAGCGTGGAGGACAAATACACACAGCTCTCTTTTGGCTTTTCCACAGCGAG 1680  
Qy 1973 CAGCACTCAGGAAGTCTGCTGCTGCGCTGAGGCTTGAAGGCTGCGACTCGGATATCTGCCC 2032  
Db 1681 CAGCACTCAGGAAGTCTGCTGCTGCGCTGAGGCTTGAAGGCTGCGACTCGGATATCTGCCC 1740  
Qy 2033 CCCAGACCTTACCCCTTCCCTGACACAGCAGCTGCTATTTTGGCACAGCTCTCTCAC 2092  
Db 1741 CCCAGACCTTACCCCTTCCCTGACACAGCAGCTGCTATTTTGGCACAGCTCTCTCAC 1800  
Qy 2093 TTCTACTGCTCAGCACTACGAGGAGCTGCGAGTCTTACTCTGCTCAGCTGCGAC 2152  
Db 1801 TTCTACTGCTCAGCACTACGAGGAGCTGCGAGTCTTACTCTGCTCAGCTGCGAC 1860  
Qy 2153 CAGCTGCCACTTGGGAGACCAAGTCTATTCTGTGCGCAGCGGCGAGCAAGCTGAC 2212  
Db 1861 CAGCTGCCACTTGGGAGACCAAGTCTATTCTGTGCGCAGCGGCGAGCAAGCTGAC 1920  
Qy 2213 AGAGCTGACTGCGCGGAGCTGGCATGAAGAGAGCCCTTTGAAAGCAGTTTAAAGCG 2272  
Db 1921 AGAGCTGACTGCGCGGAGCTGGCATGAAGAGAGCCCTTTGAAAGCAGTTTAAAGCG 1980  
Qy 2273 AGAAGCTGCCAAATGGAATTTGAGAGAGCATCTGTGAGAGACAGTCTCAGGGAAGAG 2332  
Db 1981 AGAAGCTGCCAAATGGAATTTGAGAGAGCATCTGTGAGAGACAGTCTCAGGGAAGAG 2040  
Qy 2333 CTGGGAAAGTGGGAGTCACTCAGTCTAGCTTTTCCGGCAGCATGGAAATCATGAGTCTCC 2392  
Db 2041 CTGGGAAAGTGGGAGTCACTCAGTCTAGCTTTTCCGGCAGCATGGAAATCATGAGTCTCC 2100  
Qy 2393 TGAGAGAAAGACACTGTGACTTCTATAGACAAATTTTCTTGTTCACAAAAAT 2452  
Db 2101 TGAGAGAAAGACACTGTGACTTCTATAGACAAATTTTCTTGTTCACAAAAAT 2160  
Qy 2453 TCCCTGTAAATCTGAAATATATATATGTATACATACATAT 2492  
Db 2161 TCCCTGTAAATCTGAAATATATATATGTATACATACATAT 2200

; Sequence 257, Application US/10072012  
; Publication No. US2004003493A1  
; GENERAL INFORMATION:  
; APPLICANT: Tchernev, Velizar  
; APPLICANT: Spytek, Kimberly  
; APPLICANT: Zerhusen, Bryan  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Shinkets, Richard  
; APPLICANT: Li, Li  
; APPLICANT: Gangolli, Esha  
; APPLICANT: Padigaru, Muralidhara  
; APPLICANT: Anderson, David W.  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Miller, Charles E.  
; APPLICANT: Gerlach, Valerie  
; APPLICANT: Taupier Jr, Raymond J.  
; APPLICANT: Gusev, Vladimir Y.  
; APPLICANT: Colman, Steven D.  
; APPLICANT: Wolenc, Adam R.  
; APPLICANT: Furtak, Katarzyna  
; APPLICANT: Grosse, William M.  
; APPLICANT: Alsobrook II, John P.  
; APPLICANT: Lepley, Denise M.  
; APPLICANT: Rieger, Daniel K.  
; APPLICANT: Burgess, Catherine E.  
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same  
; FILE REFERENCE: 21402-258  
; CURRENT APPLICATION NUMBER: US/10/072,012  
; CURRENT FILING DATE: 2002-01-31  
; PRIOR APPLICATION NUMBER: 60/265,102  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: 60/265,514  
; PRIOR FILING DATE: 2001-01-31  
; PRIOR APPLICATION NUMBER: 60/265,517  
; PRIOR FILING DATE: 2001-01-31  
; PRIOR APPLICATION NUMBER: 60/265,412  
; PRIOR FILING DATE: 2001-01-31  
; PRIOR APPLICATION NUMBER: 60/265,395  
; PRIOR FILING DATE: 2001-01-31  
; PRIOR APPLICATION NUMBER: 60/266,406  
; PRIOR FILING DATE: 2001-02-02  
; PRIOR APPLICATION NUMBER: 60/266,767  
; PRIOR FILING DATE: 2001-02-05  
; PRIOR APPLICATION NUMBER: 60/267,057  
; PRIOR FILING DATE: 2001-02-07  
; PRIOR APPLICATION NUMBER: 60/266,975  
; PRIOR FILING DATE: 2001-02-07  
; PRIOR APPLICATION NUMBER: 60/267,459  
; PRIOR FILING DATE: 2001-02-08  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 1391  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 257  
; LENGTH: 2071  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; US-10-072-012-257  
  
Query Match 51.4%; Score 1712.4; DB 12; Length 2071;  
Best Local Similarity 91.6%; Pred. No. 0;  
Matches 1906; Conservative 0; Mismatches 1; Indels 173; Gaps 2;  
  
Qy 502 ACCATTACATCATCTGTCGCAAAATTAAGAGAGGAGTGGGAAAGAGGACTATTGTTGTC 561  
Db 1 ACCATTACATCATCTGTCGCAAAATTAAGAGAGGAGTGGGAAAGAGGACTATTGTTGTC 60  
Qy 562 ATGGCCCATGAGATGATTGGAACCTCAATTTCTTACTGAGAGGTTGGTGGCTCTGCTGGAA 621  
Db 61 ATGGCCCATGAGATGATTGGAACCTCAATTTCTTACTGAGAGGTTGGTGGCTCTGCTGGAA 120  
Qy 622 AGTGAACGGAAGAGTCTCTCTAATTGATAGCGGCGATTTGTGGAATACATACATCC 681

Db 121 AGTGGAAACGAAAGAGTGTCTAATTGATAGCCGGCCATTGTGGAATACAAATACATCC 180  
QY 682 CACATTTTGGAGAGCCATTAATTAATCAACTGTCTCCAGCTTATGAAGCGAAGTTCCAAACAG 741  
Db 181 CACATTTTGGAGAGCCATTAATTAATCAACTGTCTCCAGCTTATGAAGCGAAGTTCCAAACAG 240  
QY 742 GACAAAGTGTAAATACAGAGTCTATCAGAGATTCAGCGAAACATAGGTTGACATTGAT 801  
Db 241 GACAAAGTGTAAATACAGAGTCTATCAGAGATTCAGCGAAACATAGGTTGACATTGAT 300  
QY 802 TGCAGTCAAGAGTTGATGTTTACGATCAAGCTCCAAAGATGTTGCTCTCTCTTCA 861  
Db 301 TGCAGTCAAGAGTTGATGTTTACGATCAAGCTCCAAAGATGTTGCTCTCTCTTCA 360  
QY 862 GACTGTTTTCTCACTGTACTCTCTGGTAACTGGAGAGAGCTTCAACTCTGTTCACTG 921  
Db 361 GACTGTTTTCTCACTGTACTCTCTGGTAACTGGAGAGAGCTTCAACTCTGTTCACTG 420  
QY 922 CTTGC----- 926  
Db 421 CTTGCAGGTGGGTTTGCTGAGTTCTCTGTTGTTTCCCTGGCCCTCTGTGAAGGAATCC 480  
QY 927 ----- 926  
Db 481 ACTTAGTCCCTACCTGCAATTTCTCAGCCTTGCTTACCTGTTGCCAAATTTGGGCCAACC 540  
QY 927 ----- AGAGCTGATG 937  
Db 541 CGAATCTTCCCAATCTTATCTTGGCTGCCAGCAGATGCTCTCAACAGAGCTGATG 600  
QY 938 CAGCAATAGGAGTTGGTTATGTTTAAATGCGAGCAATACCTGTCCAAAGCCTGACTTT 997  
Db 601 CAGCAATAGGAGTTGGTTATGTTTAAATGCGAGCAATACCTGTCCAAAGCCTGACTTT 660  
QY 998 ATCCCGAGTCTCATTCTCTGCTGCTGCTGTAATGACAGCTTTTGTGAGAAATTTTG 1057  
Db 661 ATCCCGAGTCTCATTCTCTGCTGCTGCTGTAATGACAGCTTTTGTGAGAAATTTTG 720  
QY 1058 CCGTGGTTGGACAAATCAGTAGAATTCATTGAGAAGCAAAAGCCTCCAAATGGATGTGT 1117  
Db 721 CCGTGGTTGGACAAATCAGTAGAATTCATTGAGAAGCAAAAGCCTCCAAATGGATGTGT 780  
QY 1118 CTAGTGCACTGTTAGCTGGGATCTCCGCTCCGCCACCATCGCTATCGCTACATCATG 1177  
Db 781 CTAGTGCACTGTTAGCTGGGATCTCCGCTCCGCCACCATCGCTATCGCTACATCATG 840  
QY 1178 AAGAGATGGACATGCTTTAGATCAAGCTTACAGATTGTGAAAGAAAAGAAAGACCTACT 1237  
Db 841 AAGAGATGGACATGCTTTAGATCAAGCTTACAGATTGTGAAAGAAAAGAAAGACCTACT 900  
QY 1238 ATATCTCCAACTTCAATTTTCTGGGCCAACTCTCTGGAATATGAGAAGATTAAGAAC 1297  
Db 901 ATATCTCCAACTTCAATTTTCTGGGCCAACTCTCTGGAATATGAGAAGATTAAGAAC 960  
QY 1298 CAGACTGGAGCATCAGGCGCAAGAGCAAACTCAAGCTGTGCACTTGGAGAGCCAAAT 1357  
Db 961 CAGACTGGAGCATCAGGCGCAAGAGCAAACTCAAGCTGTGCACTTGGAGAGCCAAAT 1020  
QY 1358 GAACCTGTCCTCTGCTCTCAGAGGGTGGACAGAAAAGCAGACGCCCTCAGTCCACCC 1417  
Db 1021 GAACCTGTCCTCTGCTCTCAGAGGGTGGACAGAAAAGCAGACGCCCTCAGTCCACCC 1080  
QY 1418 TGTGGCAGCTCTGCTCTCAGAGGCAGACAGAAAAGCCCGTGCATCCGCCAGCGTG 1477  
Db 1081 TGTGGCAGCTCTGCTCTCAGAGGCAGACAGAAAAGCCCGTGCATCCCGTCCG----- 1132  
QY 1478 CCGAGCGTCCCGAGCGTGCGCTGTTTAGAGGACAGCCCGCTGCTACAGCGCTC 1537  
Db 1133 -CCAGCGTACCCAGCGTGCGCTGTTTAGAGGACAGCCCGCTGCTACAGCGCTC 1191  
QY 1538 AGTGGGCTGCACCTGTCCGAGACAGGCTGGAAGACAGCAATAGCTCAAGCGTTCTTC 1597  
Db 1192 AGTGGGCTGCACCTGTCCGAGACAGGCTGGAAGACAGCAATAGCTCAAGCGTTCTTC 1251

QY 1598 TCTCTGGATATCAAAATCAGTTTCTATATTACCCAGCATGCGCAGCATCTTACATGGCTTC 1657  
Db 1252 TCTCTGGATATCAAAATCAGTTTCTATATTACCCAGCATGCGCAGCATCTTACATGGCTTC 1311  
QY 1658 TCTCTCATCAGAAAGATGCTTTGGAAATATCTACAAACCTTCCACTACTCTGGATGGACCAAC 1717  
Db 1312 TCTCTCATCAGAAAGATGCTTTGGAAATATCTACAAACCTTCCACTACTCTGGATGGACCAAC 1371  
QY 1718 AAGTATGCCAGTTCTCCCTGTTTCAGAACTATCGGAGCAGACTCCCGAAACAGTCTCT 1777  
Db 1372 AAGTATGCCAGTTCTCCCTGTTTCAGAACTATCGGAGCAGACTCCCGAAACAGTCTCT 1431  
QY 1778 GATAAGGAGGAGCCAGCATCCCAAGAGCTGCGAGACCCCGAGGCTTCAGACAGCCAG 1837  
Db 1432 GATAAGGAGGAGCCAGCATCCCAAGAGCTGCGAGACCCCGAGGCTTCAGACAGCCAG 1491  
QY 1838 AGCAAGCGATTTCATTCGGTCAGAACCCAGCAGCAGTGGCCACCGCCAGAGGTCCCTTTTA 1897  
Db 1492 AGCAAGCGATTTCATTCGGTCAGAACCCAGCAGCAGTGGCCACCGCCAGAGGTCCCTTTTA 1551  
QY 1898 TCTCCACTGCAATCGAAGTGGAGCGTGGAGACAATTACCACACAGCTTCCCTTTGGCC 1957  
Db 1552 TCTCCACTGCAATCGAAGTGGAGCGTGGAGACAATTACCACACAGCTTCCCTTTGGCC 1611  
QY 1958 CTTTCCACAGCAGCAGCACTCAGAAAGTCTGCTGGCCCTTAAAGGGCTGGCAC 2017  
Db 1612 CTTTCCACAGCAGCAGCACTCAGAAAGTCTGCTGGCCCTTAAAGGGCTGGCAC 1671  
QY 2018 TCGGATATCTTGGCCCCCAGACCTCTACCCCTTCCCTGACCCAGCAGCTGGTATTTGGCC 2077  
Db 1672 TCGGATATCTTGGCCCCCAGACCTCTACCCCTTCCCTGACCCAGCAGCTGGTATTTGGCC 1731  
QY 2078 ACAGAGTCTCTCACACTTCTACTCTGCTCAGCCATCTACCGAGCAGTGGCCAGTACTCT 2137  
Db 1732 ACAGAGTCTCTCACACTTCTACTCTGCTCAGCCATCTACCGAGCAGTGGCCAGTACTCT 1791  
QY 2138 GCCTACAGCTCAGCCAGCTGCCACTTGGGAGACCAAGTCTATTCTGTGCGAGCGG 2197  
Db 1792 GCCTACAGCTCAGCCAGCTGCCACTTGGGAGACCAAGTCTATTCTGTGCGAGCGG 1851  
QY 2198 CAGAGCCCAAGTGCAGAGCTGACTCGCGGGAGCTGGCATGAAGAGAGCCCTTTGAA 2257  
Db 1852 CAGAGCCCAAGTGCAGAGCTGACTCGCGGGAGCTGGCATGAAGAGAGCCCTTTGAA 1911  
QY 2258 AAGCAGTTTAAACCCAGAAAGCTGCCAAATGGAATTTGGAGAGAGCATCATGTGAGAAC 2317  
Db 1912 AAGCAGTTTAAACCCAGAAAGCTGCCAAATGGAATTTGGAGAGAGCATCATGTGAGAAC 1971  
QY 2318 AGGTCACGGGAGAGCTGGGAAAAGTGGGAGTGGCATCTAGCTTTTGGGCGAGATGGAA 2377  
Db 1972 AGGTCACGGGAGAGCTGGGAAAAGTGGGAGTGGCATCTAGCTTTTGGGCGAGATGGAA 2031  
QY 2378 ATCATTTAGAGTCTCTCTGAGAGAAAGACACTTTGTGACTTC 2417  
Db 2032 ATCATTTAGAGTCTCTCTGAGAGAAAGACACTTTGTGACTTC 2071

## RESULT 10

US-09-816-494-3  
; Sequence 3, Application US/09816494  
; Patent No. US20020034807A1  
; GENERAL INFORMATION:  
; APPLICANT: Meyers, Rachel A.  
; TITLE OF INVENTION: 38692 AND 21117, NOVEL DUAL SPECIFICITY  
; FILE REFERENCE: 10448-030002  
; CURRENT APPLICATION NUMBER: US/09/816,494  
; CURRENT FILING DATE: 2001-03-23  
; PRIOR APPLICATION NUMBER: US 60/191,858  
; PRIOR FILING DATE: 2000-03-24  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: FastSeq for Windows Version 4.0







;; TITLE OF INVENTION: PROGNOSIS AND HUMAN NEUROBLASTOMA WITH UNFAVORABLE PROGNOSIS  
;; FILE REFERENCE: 7388-73435  
;; CURRENT APPLICATION NUMBER: US/10/220,891  
;; PRIOR FILING DATE: 2003-03-07  
;; PRIOR APPLICATION NUMBER: JP 2000/140387  
;; PRIOR FILING DATE: 2000-05-12  
;; PRIOR APPLICATION NUMBER: JP 2000/159195  
;; PRIOR FILING DATE: 2000-03-07  
;; NUMBER OF SEQ ID NOS: 108  
;; SOFTWARE: PatentIn version 3.2  
;; SEQ ID NO 85  
;; LENGTH: 787  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; NAME/KEY: modified\_base  
;; LOCATION: (2)..(2)  
;; OTHER INFORMATION: a, t, c, g, unknown or other  
;; FEATURE:  
;; NAME/KEY: modified\_base  
;; LOCATION: (4)..(5)  
;; OTHER INFORMATION: a, t, c, g, unknown or other  
;; FEATURE:  
;; NAME/KEY: modified\_base  
;; LOCATION: (7)..(7)  
;; OTHER INFORMATION: a, t, c, g, unknown or other  
;; FEATURE:  
;; NAME/KEY: modified\_base  
;; LOCATION: (12)..(13)  
;; OTHER INFORMATION: a, t, c, g, unknown or other  
;; FEATURE:  
;; NAME/KEY: modified\_base  
;; LOCATION: (16)..(19)  
;; OTHER INFORMATION: a, t, c, g, unknown or other  
;; FEATURE:  
;; NAME/KEY: modified\_base  
;; LOCATION: (23)..(23)  
;; OTHER INFORMATION: a, t, c, g, unknown or other  
;; FEATURE:  
;; NAME/KEY: modified\_base  
;; LOCATION: (31)..(31)  
;; OTHER INFORMATION: a, t, c, g, unknown or other  
;; FEATURE:  
;; NAME/KEY: modified\_base  
;; LOCATION: (574)..(574)  
;; OTHER INFORMATION: a, t, c, g, unknown or other  
;; FEATURE:  
;; NAME/KEY: modified\_base  
;; LOCATION: (635)..(635)  
;; OTHER INFORMATION: a, t, c, g, unknown or other  
;; FEATURE:  
;; NAME/KEY: modified\_base  
;; LOCATION: (660)..(660)  
;; OTHER INFORMATION: a, t, c, g, unknown or other  
;; FEATURE:  
;; NAME/KEY: modified\_base  
;; LOCATION: (725)..(725)  
;; OTHER INFORMATION: a, t, c, g, unknown or other  
;; FEATURE:  
;; NAME/KEY: modified\_base  
;; LOCATION: (750)..(750)  
;; OTHER INFORMATION: a, t, c, g, unknown or other  
;; FEATURE:  
;; NAME/KEY: modified\_base  
;; LOCATION: (752)..(752)  
;; OTHER INFORMATION: a, t, c, g, unknown or other  
;; FEATURE:  
;; NAME/KEY: modified\_base  
;; LOCATION: (755)..(755)  
;; OTHER INFORMATION: a, t, c, g, unknown or other  
;; FEATURE:  
;; NAME/KEY: modified\_base  
;; LOCATION: (760)..(760)

;; OTHER INFORMATION: a, t, c, g, unknown or other  
;; FEATURE:  
;; NAME/KEY: modified\_base  
;; LOCATION: (763)..(763)  
;; OTHER INFORMATION: a, t, c, g, unknown or other  
;; FEATURE:  
;; NAME/KEY: modified\_base  
;; LOCATION: (765)..(765)  
;; OTHER INFORMATION: a, t, c, g, unknown or other  
;; FEATURE:  
;; NAME/KEY: modified\_base  
;; LOCATION: (771)..(771)  
;; OTHER INFORMATION: a, t, c, g, unknown or other  
;; FEATURE:  
;; NAME/KEY: modified\_base  
;; LOCATION: (777)..(777)  
;; OTHER INFORMATION: a, t, c, g, unknown or other  
US-10-220-891-85  
Query Match 17.9%; Score 595; DB 15; Length 787;  
Best Local Similarity 96.6%; Pred. No. 4.6e-175; Indels 5; Gaps 5;  
Matches 659; Conservative 0; Mismatches 18;  
QY 2651 TTCAGTGAATTACACATCCTCTTCTTCTTAAAGCAAGTGTCTTTGGTGGAGGAC 2710  
DB 45 TACTGGGAATTACACATCCTCTTCTTCTTAAAGCAAGTGTCTTTGGTGGAGGAC 104  
QY 2711 AAATCCCTACCAATTTTCCACGTTGTCTACTAAGAGATCTCAATATATTAGTCTTTGTC 2770  
DB 105 AAATCCCTACCAATTTT-CACGTTGTCTACTAAGAGATCTCAATATATTAGTCTTTGTC 163  
QY 2771 CGGACCCCTTCCATAGTACACCTTAGCGCTGAGACTGAGCCAGCTTGGGGGTCAAGTAGGT 2830  
DB 164 CGGACCCCTTCCATAGTACACCTTAGCGCTGAGACTGAGCCAGCTTGGGGGTCAAGTAGGT 223  
QY 2831 AGACCCCTTAGGGACAGAGCGCTAGTGTGTAAATCCAAAGAAATGATCTCTATCAAGCT 2890  
DB 224 AGACCCCTTAGGGACAGAGCGCTAGTGTGTAAATCCAAAGAAATGATCTCTATCAAGCT 283  
QY 2891 GATTCAAAACCCAGCTCACCCTGACAGCGGAGGACAGCATCACTCTGCTGGACGG 2950  
DB 284 GATTCAAAACCCAGCTCACCCTGACAGCGGAGGACAGCATCACTCTGCTGGACGG 343  
QY 2951 ACCATTAGGGCCCTTGCCAAAGGTCTACCTTTAGAGCAAAACCCAGTACCTCAGAGGAAAG 3010  
DB 344 ACCATTAGGGCCCTTGCCAAAGGTCTACCTTTAGAGCAAAACCCAGTACCTCAGAGGAAAG 403  
QY 3011 TCGGGCTTGACCACTACCATATCTGTAGCCCATTTCTAGGCATTTGTGAATAGTAG 3070  
DB 404 TC-GGGCTTTGACCACTACCATATCTGTAGCCCATTTCTAGGCATTTGTGAATAGTAG 462  
QY 3071 GTAGCTAGTCACACTTTTTCAGACCAATTCAAACTGTCTATGCAAAATTCCTGCGGCC 3130  
DB 463 GTAGCTAGTCACACTTTTTCAGACCAATTCAAACTGTCTATGCAAAATTCCTGCGGCC 521  
QY 3131 TAGATGGAGATAATTTTTTTTCTCAGCTTTATGAAGAGGGAAGAACTGTCTAGGA 3190  
DB 522 TAGATGGAGATAA-TTTTTTTTCTCAGCTTTATGAAGAGGGAAGAACTGTCTAGGA 580  
QY 3191 TTCAGCTCAACCAACAGGAACCTGGCAACATTCAGATTTAAGCTTAAGTTGGGAGGCTAA 3250  
DB 581 TTCAGCTCAACCAACAGGAACCTGGCAACATTCAGATTTAAGCTTAAGTTGGGAGGCTAA 640  
QY 3251 CGAGTCTACCTCCCTCTTTTGTAAATCAAGAAATGTTTTAAATGGGATTTGTAATCCTTT 3310  
DB 641 CGAGTCTACT-CCCTCTTGTAAATCAAGAAATGTTTTAAATGGGATTTGTAATCCTTT 699  
QY 3311 AAATAAGATGAATCTGTTTC 3332  
DB 700 AAATAAGATGAATCTGTTTC 721  
RESULT 14

```

US-09-764-853-158
; Sequence 158, Application US/09764853
; Patent No. US20020090672A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, P
; FILE REFERENCE: P3206
; CURRENT APPLICATION NUMBER: US/09/76
; PRIOR FILING DATE: 2001-01-17
; Prior application data removed - con
; NUMBER OF SEQ ID NOS: 939
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 158
; LENGTH: 877
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-853-158

```

Query Match	16.1%;	Score 537.4;	DB 9;	Length 877;
Best Local Similarity	98.5%;	Pred. No. 6e-157;		
Matches 574;	Conservative	0;	Mismatches 6;	Indels 3;
QY	285	GTGGAATGGTTATTCACGTCATCTCTTATGAATCAAAATGTGAGGGGCGCTTTGTGGAC	344	
DB	15	GTGGAATGGTTATTCACGTCATCTCTTATGAATCAAAATGTGAGGGGCGCTTTGTGGAC	74	
QY	345	GGAGTCCITTTGCAAGAGCACAATCAACGGGAAGGAAGAGACATTTCATCTGGAGGGCTC	404	
DB	75	GGAGT-CTTTGCAAGAGCACAATCAACGGGAAGGAAGAGACATTTCATCTGGAGGGCTC	133	
QY	405	TTTGCGAAAATGGGTTTAACTCTCTTTGCCAGTCACCAACGAGCTGACCTCATACACT	464	
DB	134	TTTGCGAAAATGGGTTTAACTCTCTTTGCCAGTCACCAACGAGCTGACCTCATACACT	193	
QY	465	TTTATGTACAAATGGAGTGGCTCAGCGCTTTTGAGCAACACCAACATTACATCATCTGGTGC	524	
DB	194	TTTATGTACAAATGGAGTGGCTCAGCGCTTTTGAGCAACACCAACATTACATCATCTGGTGC	253	
QY	525	TAAAGAGGAGGTGGGAAAAGAGGACTATTGTGTGTCATGGCCCATGAGATGATTTGGAAC	584	
DB	254	TAAAGAGGAGGTGGGAAAAGAGGACTATTGTGTGTCATGGCCCATGAGATGATTTGGAAC	313	
QY	585	TCAAATTTGTTACTGAGAGGTTGGTGGCTCTGCTGGAAAGTGGAAACGGAAAAAGTGTGCT	644	
DB	314	TCAAATTTGTTACTGAGAGGTTGGTGGCTCTGCTGGAAAGTGGAAACGGAAAAAGTGTGCT	373	
QY	645	AATTGATAGCGGCGCATTTGTGGAAATACATACATCCACATTTTGGAGGCCATTAATAT	704	
DB	374	AATTGATAGCGGCGCATTTGTGGAAATACATACATCCACATTTTGGAGGCCATTAATAT	433	
QY	705	CAACTGCTCCAGGCTTATGAAGCGAAGGTTGCACAGGACAAAGTGTTAATTACAGAGCT	764	
DB	434	CAACTGCTCCAGGCTTATGAAGCGAA- GTTGCACAGGACAAAGTGTTAATTACAGAGCT	492	
QY	765	CATCCAGCATTCAGCGAACCATAGGTTGCATGATTGTCAGTCCAGAGGTTGTAGTTTA	824	
DB	493	CATCCAGCATTCAGCGAACCATAGGTTGCATGATTGCA-TCAGAGGTTGTAGTTTA	551	
QY	825	CGATCAAAGCTCCCAAGATGTTGGCTCTCTCTCTTCAGACTGT	867	
DB	552	CGATCAAAGCTCCCAAGATGTTGGCTGTCCTCTCTTCAGCATGT	594	

RESULT 15

```

RESOL1 15
US-09-918-995-25801
; Sequence 25801, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995

```

```

; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: Fast-SEQ for Windows Version 3.0
; SEQ ID NO 25801
; LENGTH: 478
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(478)
; OTHER INFORMATION: n = A,T,C or G
; US-09-918-995-25801

```

Query Match	12.3%; Score 409.2; DB 10; Length 478;
Best Local Similarity	98.6%; Pred. No. 6.4e-117;
Matches 433; Conservative	0; Mismatches 4; Indels 2; Gaps 2;
QY	2499 GGAAAAATGGAGCTATGGTGTAAAGGCAACAGGTGGATCAACCCAGTTGTTACTCTCTCTAA 2558
DB	42 GGAAAAATGGAGCTATGGTGTAAAGGCAACAGGTGGATCAACCCAGTTGTTACTCTCTCTAA 101
QY	2559 CATCTGCAATTTGAGAGATCAGCTAATPACTTCTCTCAACAAAAATGGAAGGCGAGATCCTA 2618
DB	102 CATCTGCAATTTGAGAGATCAGCGAATPACTTCTCTCAACAAAAATGGAAGGCGAGATCCTA 161
QY	2619 GAATCCCCCTTAGACGAGGAGAAACCAATTTATTTCAGTGAATTACACATCCTCTTTGTTCT 2678
DB	162 GAATCCCCCTTAGACGAGGAGAAACCAATTTATTTCAGTGAATTACACATCCTCTTTGTTCT 221
QY	2679 TAAAAAGCAAGTGCTCTTTGGTGTGGAGGACAAAAATGCCCTACCATTTCCAGTTGTG 2738
DB	222 TAAAAAGCAAGTGCTCTTTGGTGTGGAGGACAAAAATGCCCTACCATTTT-CACGTTGTG 280
QY	2739 CTACTTAGAGATCTCAAAATATTAGTCTTTGTCGGACGCTTCCATAGTACACCTTTAGCGC 2798
DB	281 CTACTTAGAGATCTCAAAATATTAGTCTTTGTCGGACGCTTCCATAGTACACCTTTAGCGC 340
QY	2799 TGAGACTGAGCCAGCTTGGGGGTACGGTAGGTAGACCTGTTTAGGACAGAGCCTTAGTGG 2858
DB	341 TGAGACTGAGCCAGCTTGGGGGTACGGTAGGTAGACCTGTTTAGGACAGAGCCTTAGTGG 400
QY	2859 TAAATCCAAGAGAAATGATCCTATCAAAAGCTGATTTCAAAACCCACGCTCACTGTGACAG 2918
DB	401 -AAATCCAAGAGAAATGATCCTATCCAAAGCTGATTTCAAAACACACGCTCACTGTGACAG 459
QY	2919 CCGAGGGACAGGATCA 2937
DB	460 NCGAGGACAGGATCA 478

Search completed: February 28, 2004, 01:16:00  
Job time : 1085 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 27, 2004, 18:43:34 ; Search time 8123 Seconds  
(without alignments)  
12249.266 Million cell updates/sec

Title: US-09-964-277-20

Perfect score: 3332

Sequence: 1 gagagaaagagaataata.....ataaagatgaacttggttc 3332

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*

2: em\_esthm:\*

3: em\_estin:\*

4: em\_estmu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_htc:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_htc:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estom:\*

17: em\_gss\_hum:\*

18: em\_gss\_inv:\*

19: em\_gss\_pln:\*

20: em\_gss\_vrt:\*

21: em\_gss\_fun:\*

22: em\_gss\_mam:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_rtd:\*

26: em\_gss\_phg:\*

27: em\_gss\_vrl:\*

28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	896.4	26.9	1201	13	BX332176
2	858	25.8	1060	12	BM546940
3	834.6	25.0	883	13	BQ945892
4	827.8	24.8	881	13	BQ930140

5	809.8	24.3	913	13	BUL70187
6	809.6	24.3	859	13	BQ933499
7	786.2	23.6	974	13	BQ951695
8	777	23.3	798	13	EX110818
9	766.4	23.0	903	12	BG482429
10	741.2	22.2	891	13	BQ670989
11	736.4	22.1	1148	12	BM906608
12	717.8	21.5	740	10	AW963248
13	688.4	20.7	720	14	CA944683
14	641.2	19.2	920	10	BE897795
15	632	19.0	663	13	EX479029
16	615.4	18.5	660	12	BI917706
17	607.4	18.2	3325	11	AK035652
18	606.6	18.2	983	12	BG155198
19	603.2	18.1	752	13	BQ432082
20	574.4	17.2	650	9	AU253268
21	574	17.2	593	12	BM930580
22	551.8	16.6	685	14	CB165612
23	551.4	16.5	557	9	AI807619
24	541.2	16.2	621	9	AV703072
25	539	16.2	539	9	AI659120
26	537	16.1	548	9	AI637845
27	537	16.1	554	9	AW014773
28	536.8	16.1	693	9	AV701628
29	534.8	16.1	543	9	AI936544
30	534.8	16.1	790	12	BI661614
31	533	16.0	551	9	AI674392
32	511	15.3	723	13	BU708175
33	508.8	15.3	524	10	BF432802
34	493	14.8	496	9	AI678804
35	492	14.8	862	13	BQ770036
36	490.8	14.7	743	14	CF727802
37	485	14.6	796	10	BF983182
38	484.6	14.5	514	12	BM684128
39	483.4	14.5	524	10	BF758523
40	482	14.5	511	12	BM930518
41	479.4	14.4	739	28	AZ850283
42	466.6	14.0	546	10	BE232856
43	465	14.0	898	13	BQ721265
44	462.4	13.9	631	28	BH039241
45	462.4	13.9	1067	10	BF135687

ALIGNMENTS

RESULT 1  
BX332176  
LOCUS  
DEFINITION BX332176 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens  
CDNA clone CS0DC008YN23 5-PRIME, mRNA sequence.  
ACCESSION BX332176  
VERSION BX332176.1 GI:30335119  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 1201)  
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT Genoscope - Centre National de Sequençage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 6659.r For  
more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CS0DC008CG12QPI&cluster=6659.r. Contact :  
Feng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600

1201 bp mRNA linear EST 02-MAY-2003  
BX332176 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens  
CDNA clone CS0DC008YN23 5-PRIME, mRNA sequence.  
BX332176  
BX332176.1 GI:30335119  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1201)  
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished (2001)  
Genoscope - Centre National de Sequençage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 6659.r For  
more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CS0DC008CG12QPI&cluster=6659.r. Contact :  
Feng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600



```
QY 1107 ATGGATGTCTTCTAGTGCACTGTTTAGCTGGAGTCTCCGCTCCGCCACCATCGCTATCG 1166
Db 229 ATGGATGTCTTCTAGTGCACTGTTTAGCTGGAGTCTCCGCTCCGCCACCATCGCTATCG 288
QY 1167 CTTACATCATGAAGAGGATGGACATGTCTTTAGATGAAGCTTACAGATTGTGAAAGAAA 1226
Db 289 CTTACATCATGAAGAGGATGGACATGTCTTTAGATGAAGCTTACAGATTGTGAAAGAAA 348
QY 1227 AAAGACCTACTATCTCCAACTTCAATTTCTGGGCCACTCTCTGGACTATGAGAGA 1286
Db 349 AAAGACCTACTATCTCCAACTTCAATTTCTGGGCCACTCTCTGGACTATGAGAGA 408
QY 1287 AGATTAAAGAACAGACTGGAGCATCAGGCGCCAAAGAGCAAACTCAAGCTGTGCACCTGG 1346
Db 409 AGATTAAAGAACAGACTGGAGCATCAGGCGCCAAAGAGCAAACTCAAGCTGTGCACCTGG 468
QY 1347 AGAAGCCAAATCAACTGTCTCCCTGCTGTCTCAGAGGTGGACAGAAAGCGAGAGCGCCC 1406
Db 469 AGAAGCCAAATCAACTGTCTCCCTGCTGTCTCAGAGGTGGACAGAAAGCGAGAGCGCCC 528
QY 1407 TCAGTCCACCCCTGTGCCACTCTGCTACTCTCAGAGGCGAGGACAAAGGCCGTGCATC 1466
Db 529 TCAGTCCACCCCTGTGCCACTCTGCTACTCTCAGAGGCGAGGACAAAGGCCGTGCATC 588
QY 1467 CGCCAGCGTGCAGCGTGTGCCAGCGTGCAGCGGTGCTGTTAGAGGACAGCCCGCTGG 1526
Db 589 CGC-----CCAGCGTACCCAGCGTGCAGCGGTGCTGTTAGAGGACAGCCCGCTGG 639
QY 1527 TACAGCGCTCAGTGGCTGCACCTGTCGCGAGCAGCGCTGGAAGACAGCAATAAGCTCA 1586
Db 640 TACAGCGCTCAGTGGCTGCACCTGTCGCGAGCAGCGCTGGAAGACAGCAATAAGCTCA 699
QY 1587 AGCGTTCCCTTCTCTCTGGATATCAATCAGTTTCAATTCAGCCAGCATGGCAGCATCCT 1646
Db 700 AGCGTTCCCTTCTCTCTGGATATCAATCAGTTTCAATTCAGCCAGCATGGCAGCATCCT 759
QY 1647 TACATGGCTTCTCTCATCAGAGATGCTTTGGATACTACTACAACTTCCACTACTCTGG 1706
Db 760 TACATGGCTTCTCTCATCAGAGATGCTTTGGATACTACTACAACTTCCACTACTCTGG 819
QY 1707 ATGGGACCAACAAGCTATGTCAGATTTCTCCCTGTTCAGGAACTATCGGAGCAGACTCCCG 1766
Db 820 ATGGGACCAACAAGCTATGTCAGATTTCTCCCTGTTCAGGAACTATCGGAGCAGACTCCCG 879
QY 1767 AAACGAGTCTGATTAAGAGAGAGCGCAGCATCCCGAAGAGCTGCGACCGCCA-GGCTT 1825
Db 880 AAACGAGT-CTGATAAGAGAGAGCGCAGCATCCCGAAGAGCTGCGACCGCGCGGCTT 938
QY 1826 TCAGACAGCCAGCAAGCG-----ATTGCATTGCTGATGCAATCCAGCCAGCAGTGGCAAC 1879
Db 939 TCAGACAGCCAGCAAGCGGATGCAATCCAGCCAGCAGTGGCAAC 1904
QY 1880 GCCAGAGGTCCTTTTATCTCCAC 1904
Db 999 CCAGAAAGTCCCTTTTATCTCC 1023

RESULT 3
BQ945892
LOCUS
DEFINITION BQ945892 883 bp mRNA linear EST 21-AUG-2002
5', mRNA sequence.
ACCESSION BQ945892
VERSION BQ945892.1 GI:22361370
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 883)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
```

## COMMENT

Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-remail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone Distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLCM2653 row: o column: 14  
High quality sequence stop: 672.

## FEATURES

source  
1. 883  
location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6463045"  
/tissue\_type="epidermoid carcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC 101"

/note="Organ: lung; Vector: pOTB7; Site 1: EcoRI; Site 2:  
XhoI; cDNA made by oligo-dT priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:  
GGCAGCAG(G). Library constructed by Ling Hong in the  
Laboratory of Gerald M. Rubin (University of California,  
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
Superscript II RT (Life Technologies). Note: this is a  
NIH\_MGC Library."

## ORIGIN

Query Match 25.0%; Score 834.6; DB 13; Length 883;  
Best Local Similarity 98.5%; Pred. No. 3.9e-215;  
Matches 874; Conservative 0; Mismatches 9; Indels 4; Gaps 3;  
QY 1191 TGCTTTAGATGAAGCTTACAGATTGTTGGAAGAAAGAACCTACTATATCTCCAACT 1250  
Db 1 TGCTTTAGATGAAGCTTACAGATTGTTGGAAGAAAGAAAGA-CTACTATATCTCCAA--C 57  
QY 1251 TCAATTTTCTGGGCCAACTCCTGGACTATGAGAAGAGATTAAAGAACCCAGACTGGAGCAT 1310  
Db 58 TTCAATTTTCTGGGCCAACTCCTGGACTATGAGAAGAGATTAAAGAACCCAGACTGGAGCAT 117  
QY 1311 CAGGCCCAAGAGCAAACTCAAGCTGTGCACTGGAGAGCCAAATGACCTGTCCCTG 1370  
Db 118 CAGGCCCAAGAGCAAACTCAAGCTGTGCACTGGAGAGCCAAATGACCTGTCCCTG 177  
QY 1371 CTGTCTCAGAGGTGGACAGAAAGCGAGACGCGCCCTCAGTCCACCCCTGTGCCGACTCTG 1430  
Db 178 CTGTCTCAGAGGTGGACAGAAAGCGAGACGCGCCCTCAGTCCACCCCTGTGCCGACTCTG 237  
QY 1431 CTACTCAGAGCGAGGACAGCAAAAGGCCCTGTGATCCCGCAGCGTGCAGCGGCCCA 1490  
Db 238 CTACTCAGAGCGAGGACAGCAAAAGGCCCTGTGATCCCGCAGCGTGCAGCGGCCCA 297  
QY 1491 CGGTGACCGCTGCTGTTAGAGGACAGCCGCTGTGACAGGCGCTCAGTGGGCTGCACC 1550  
Db 298 CGGTGACCGCTGCTGTTAGAGGACAGCCGCTGTGACAGGCGCTCAGTGGGCTGCACC 357  
QY 1551 TGTCGCGAGCAGGCTGGAAAGACAGCAATAAGCTCAAGCGTTCTCTCTCTGGATATCA 1610  
Db 358 TGTCGCGAGCAGGCTGGAAAGACAGCAATAAGCTCAAGCGTTCTCTCTCTGGATATCA 417  
QY 1611 AATCAGTTTCAATTCAGCCAGCATGGCAGCATCCTTACATGGCTTCTCTCTCATCAGAAG 1670  
Db 418 AATCAGTTTCAATTCAGCCAGCATGGCAGCATCCTTACATGGCTTCTCTCTCATCAGAAG 477  
QY 1671 ATGCTTTGGATATCAAAACCTTCCACTACTCTGGATGGAGCAACCAAGCTATGCCAGT 1730  
Db 478 ATGCTTTGGATATCAAAACCTTCCACTACTCTGGATGGAGCAACCAAGCTATGCCAGT 537  
QY 1731 TCTCCCTGTTTCAGGAACATATCGGAGCAGACTCCCGAAACCACTCTGTATAGAGGAGAG 1790  
Db 538 TCTCCCTGTTTCAGGAACATATCGGAGCAGACTCCCGAAACCACTCTGTATAGAGGAGAG 597

```
QY 1791 CCAGCATCCCAAGAGCTGCAGACCGCCAGCCCTTCAGACGCCAGAGGAGGATTC 1850
D 598 CCAGCATCCCAAGAGCTGCAGACCGCCAGCCCTTCAGACGCCAGAGGAGGATTC 657
QY 1851 ATTTCGGTTCAGAACCCAGCAGCAGTGGCACCAGCCAGAGGTCCTTTTATCTCCACTGCATC 1910
D 658 ATTTCGGTTCAGAACCCAGCAGCAGTGGCACCAGCCAGAGGTCCTTTTATCTCCACTGCATC 717
QY 1911 GAAGTGGGAGCGTGGAGGACAATTACCAACAGAGCTTCCTTTTCGGCCCTTTCCACAGCC 1970
D 718 GAAGTGGGAGCGTGGAGGACAATTACCAACAGAGCTTCCTTTTCGGCCCTTTCCACAGCC 777
QY 1971 AGCAGCACTTCACAACTTCCTGGCTGGGCTTAAAGGCTGGCAGCTGCATATCTGG 2030
D 778 AGCAGCACTTCACAACTTCCTGGCTGGGCTTAAAGGCTGGCAGCTGCATATCTGG 836
QY 2031 CCCCCAGACCTTACCCCTTCCTGACAGCAGTGGTATTTGGCC 2077
D 837 CCCCCAGACCTTACCCCTTCCTGACAGCAGTGGTATTTGGCC 883

RESULT 4
BQ930140
LOCUS
DEFINITION BQ930140 881 bp mRNA linear EST 20-AUG-2002
5', mRNA sequence.
AGENCOURT.8923732 NIH_MGC_101 Homo sapiens cdna clone IMAGE:6460964
ACCESSION BQ930140
VERSION BQ930140.1 GI:22345171
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 881)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2648 row: h column: 21
High quality sequence stop: 676.
FEATURES
Location/Qualifiers
1..881
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6460964"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="PH10B (phage-resistant)"
/clone_lib="NIH_MGC_101"
/notes="Organ: lung; Vector: pOTB7; Site:1: EcoRI; Site 2:
XhoI; cDNA made by oligo-dt priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

ORIGIN
Query Match 24.8%; Score 827.8; DB 13; Length 881;
Best Local Similarity 99.4%; Pred. No. 2.8e-213;
Matches 862; Conservative 0; Mismatches 2; Indels 3; Gaps 3;
QY 1191 TGCTTTAGTAGAGCTTACAGATTGTGAAAGAAAAAGACCTACTATATCTCAAAC 1250
```

```
Db 1 TGTCTTTAGTAGAGCTTACAGATTGTGAAAGAAAAAGACCTACTATATCTCAAAC 60
QY 1251 TCAATTTTCTGGGCCAACTCCTCGACTATGAGAGAGATTAAAGAACAGACTGGAGCAT 1310
D 61 TCAATTTTCTGGGCCAACTCCTCGACTATGAGAGAGATTAAAGAACAGACTGGAGCAT 120
QY 1311 CAGGCCCAAGAGCAAACTCAAGCTGCTGCACCTGGAGAGAGCAAAATGAACCTTCCCTG 1370
D 121 CAGGCCCAAGAGCAAACTCAAGCTGCTGCACCTGGAGAGAGCAAAATGAACCTTCCCTG 180
QY 1371 CTGCTCTCAGAGGGTGGACAGAAAGCCAGACGCCCTCAGTCCACCCCTGTGCCACTCTG 1430
D 181 CTGCTCTCAGAGGGTGGACAGAAAGCCAGACGCCCTCAGTCCACCCCTGTGCCACTCTG 240
QY 1431 CTACCTCAGAGGCAGCAGGACAAAGGCCCGTGCATCCGCCAGCGTGCACAGGTGCCCA 1490
D 241 CTACCTCAGAGGCAGCAGGACAAAGGCCCGTGCATCCGCCAGCGTGCACAGGTGCCCA 300
QY 1491 GCGTGCAGCCGCTGCTGTAGAGGACAGCCCGCTGTACAGGCGCTCAGTGGGCTGCACC 1550
D 301 GCGTGCAGCCGCTGCTGTAGAGGACAGCCCGCTGTACAGGCGCTCAGTGGGCTGCACC 360
QY 1551 TGTCCGACAGAGCTGGAAGACAGCAATAGCTCAAGCGTTCTTCTCTGGATATCA 1610
D 361 TGTCCGACAGAGCTGGAAGACAGCAATAGCTCAAGCGTTCTTCTCTGGATATCA 420
QY 1611 AATCAGTTTCTATATTTCAGCCAGCAGTGCAGCATCTTACATGGCTTCTCTCATCAGAAG 1670
D 421 AATCAGTTTCTATATTTCAGCCAGCAGTGCAGCATCTTACATGGCTTCTCTCATCAGAAG 480
QY 1671 ATGCTTTTGAATACTACAACTTTCCTACTCTTGGATGGAGCAACAAGCTATGCCAGT 1730
D 481 ATGCTTTTGAATACTACAACTTTCCTACTCTTGGATGGAGCAACAAGCTATGCCAGT 540
QY 1731 TCTCCCTGTTTCAGAACTATCGAGCAGACTCCGAAACCGCTCTCTGATAAGAGGAG 1790
D 541 TCTCCCTGTTTCAGAACTATCGAGCAGACTCCGAAACCGCTCTCTGATAAGAGGAG 600
QY 1791 CCAGCATCCCCAAGAAAGCTGCAGACCGCCAGCCCTTCAGACAGCCAGAGCGATTGC 1850
D 601 CCAGCATCCCCAAGAAAGCTGCAGACCGCCAGCCCTTCAGACAGCCAGAGCGATTGC 660
QY 1851 ATTTCGTTCAGAACCCAGCAGCAGTGGCAGCCCGCCAGAGTCTCTTATCTCCACTGCATC 1910
D 661 ATTTCGTTCAGAACCCAGCAGCAGTGGCAGCCCGCCAGAGTCTCTTATCTCCACTGCATC 720
QY 1911 GAAGTGGGAGCGTGGAGGACAATTACCAACAGCTTCC-TTTTCGGCCTTTCCACCAGC 1969
D 721 GAAGTGGGAGCGTGGAGGACAATTACCAACAGCTTCC-TTTTCGGCCTTTCCACCAGC 780
QY 1970 CAGCAGCACCTCAGAACTGTGTGGCT-GGGCTTTAAGGGCTGGCAGCTCGGATATCTT 2028
D 781 CAGCAGCACCTCAGAACTGTGTGGCTGGGCTTTAAGGGCTGGCAGCTCGGATATCTT 840
QY 2029 GG-CCCCCCCAGACCTCTACCCCTTCCC 2054
D 841 GGCCCCCCCAGACCTCTACCCCTTCCC 867

RESULT 5
BUI70187
LOCUS BUI70187
DEFINITION AGENCOURT.7940023 NIH_MGC_68 Homo sapiens cdna clone IMAGE:6007510
5', mRNA sequence.
ACCESSION BUI70187
VERSION BUI70187.1 GI:22684171
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 913)
```

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.

Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
Tissue Procurement: DCTD/BTP/Gaadar  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone Distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>

Plate: L1AM13190 Row: n column: 23  
High quality sequence stop: 643.

## FEATURES

Location/Qualifiers  
1..913  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clones="IMAGE:6007510"  
/tissue\_type="large cell carcinoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH MGC\_68"  
/note="Organ: lung; Vector: pCMV-SPORT6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 1.8 kb. Library constructed by Life  
Technologies."

## ORIGIN

Query Match 24.3%; Score 809.8; DB 13; Length 913;  
Best Local Similarity 96.3%; Pred. No. 2.2e-208;  
Matches 869; Conservative 0; Mismatches 22; Indels 6; Gaps 4;  
QY 2160 CCACCTTGGGAGACCAAGTCTATTCTGTGCGCGGCGGAGAGCCAGTGCAGAGCTG 2219  
DB 5 CCACCTTGGGAGACCAAGTCTATTCTGTGCGCGGCGGAGAGCCAGTGCAGAGCTG 64  
QY 2220 ACTCGCGCGGAGCTGGCATGAAGAGAGCCCTTTTGAAGACGAGTCTTAAACGAGAGCT 2279  
DB 65 ACTCGCGCGGAGCTGGCATGAAGAGAGCCCTTTTGAAGACGAGTCTTAAACGAGAGCT 124  
QY 2280 GCCAAATGGAATTTGGAGAGGAGCATATGTCAGAGACAGTCCAGGAGAGCTGGGA 2339  
DB 125 GCCAAATGGAATTTGGAGAGGAGCATATGTCAGAGACAGTCCAGGAGAGCTGGGA 184  
QY 2340 AAGTGGGCGAGTCAGTCTAGCTTTTTCGCGGAGCATGGAATCATTTAGGTCCTCTGAGAG 2399  
DB 185 AAGTGGGCGAGTCAGTCTAGCTTTTTCGCGGAGCATGGAATCATTTAGGTCCTCTGAGAG 244  
QY 2400 AAAGACACTTGTGACTCTATAGACAAATTTTTTTTCTTTTTCACAAAAAATTCCTCT 2459  
DB 245 AAAGACACTTGTGACTCTATAGACAAATTTTTTTTCTTTTTCACAAAAAATTCCTCT 304  
QY 2460 AAATCTGAATATATATATGTCATACATATATATTTTGGAAATGGAGCTATGGTGA 2519  
DB 305 AAATCTGAATATATATATGTCATACATATATATTTTGGAAATGGAGCTATGGTGA 364  
QY 2520 AAAGCAACAGGTGGATCAACCCAGTTTACTCTCTTAAATCTGATTTGAGAGATCAG 2579  
DB 365 AAAGCAACAGGTGGATCAACCCAGTTTACTCTCTTAAATCTGATTTGAGAGATCAG 424  
QY 2580 CTAATACCTCTCTCAACAAAATGGAAGGAGGAGATGTAATCCCTCCCTAGACGAGGA 2639  
DB 425 CTAATACCTCTCTCAACAAAATGGAAGGAGGAGATGTAATCCCTCCCTAGACGAGGA 484  
QY 2640 AAACCAATTTATTCAGTGAATTAACATCTCTTGTCTTTTAAAGCAAGTGTCTTTGG 2699  
DB 485 AAACCAATTTATTCAGTGAATTAACATCTCTTGTCTTTTAAAGCAAGTGTCTTTGG 544  
QY 2700 TGTGGAGGACAAAATCCCTTACCATTTTCCAGTTGTGCTACTAAGAGATCTCAAAAT 2759  
DB 545 TGTGGAGGACAAAATCCCTTACCATTTT-CAGTTGTGCTACTAAGAGATCTCAAAAT 603

QY 2760 TAGCTTTGTCCGGACCCCTTCCATAGTACACCTTAGCGTGAAGCTGAGCCAGCTTTGGG 2819  
DB 604 TAGCTTTGTCCGGACCCCTTCCATAGTACACCTTAGCGTGAAGCTGAGCCAGCTTTGGG 663  
QY 2820 GTGAGGTAGGTAGACCTGTAGGACAGAGCCTAGTGTAAAT-CCAGAGAAATGATC 2878  
DB 664 GTGAGGTAGGTAGACCTGTAGGACAGAGCCTAGTGTAAATCCAGAGAAATGATC 723  
QY 2879 CTATCCAAAGCTGATTTCACAAACCCAGCTCACCTGACAGCCGAGGACACGAGCATCAC 2938  
DB 724 CTATCCAAAGCTGATTTCACAAACCCAGCTCACCTGACAGCCGAGGACACGAGCATCAC 783  
QY 2939 TCTGCTGGACGGACCATTAGGGCCCTTCCCAAGG-CTTACCTTAGAGCAAAACCCAGTA 2995  
DB 784 TCTGCTGGACGGACCATTAGGGCCCTTCCCAAGGCTTACCTTAAAGCCAAACCCAGTA 843  
QY 2996 CCTCAGACAGAGAAAGT-CGGGGCTTTGACCACTACCATATCTGTAGCCCATTTTCT 3051  
DB 844 CCTCAGACAGAGAAAGTCCGGGCTTTGACCACTACCATATCTGTAGCCCATTTTCT 900

## RESULT 6

BQ933499  
LOCUS BQ933499 859 bp mRNA linear EST 21-AUG-2002  
DEFINITION AGENCOURT 8732489 NIH\_MGC\_101 Homo sapiens cDNA clone IMAGE:6455595  
5', mRNA sequence.

ACCESSION BQ933499

VERSION BQ933499.1 GI:22348892

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 859)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

JOURNAL

COMMENT

Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)

Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone Distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LICM2634 Row: i Column: 04

High quality sequence stop: 727.

## FEATURES

Location/Qualifiers  
1..859  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clones="IMAGE:6455595"  
/tissue\_type="epidermoid carcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH MGC\_101"  
/notes="Organ: lung; Vector: pOTB7; Site 1: EcoRI; Site 2:  
XhoI; cDNA made by oligo-dT priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:  
GGCACAG(G) Library constructed by Ling Hong in the  
Laboratory of Gerald M. Rubin (University of California,  
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
Superscript II RT (Life Technologies). Note: this is a  
NIH\_MGC Library."

## ORIGIN

Query Match 24.3%; Score 809.8; DB 13; Length 859;  
Best Local Similarity 96.3%; Pred. No. 2.5e-208;  
Matches 827; Conservative 0; Mismatches 11; Indels 2; Gaps 1;  
QY 1191 TGCTTTTAGTGAAGCTTACAGATTGTGAAAGAAAAAGACCTACTATCTCTCAAACT 1250

Db 1 TGTCTTAGATGAAGCTTACAGATTGTGGAAGAAAAAGACCTACTATATCTCCAACT 60  
Qy 1251 TCAATTTCTGGCCAACTCTGGACTATGAGAGAGATTAAAGACAGACTGGAGCAT 1310  
Db 61 TCAATTTCTGGCCAACTCTGGACTATGAGAGAGATTAAAGACAGACTGGAGCAT 120  
Qy 1311 CAGGGCCAAAGAGCAAACTCAAGCTGCTGCACTCTGGAGAACCAAAATGAACCTGTCCCTG 1370  
Db 121 CAGGGCCAAAGAGCAAACTCAAGCTGCTGCACTCTGGAGAACCAAAATGAACCTGTCCCTG 180  
Qy 1371 CTGTCTCAGAGGTTGGACAGAAAAGAGAGAGCCCTCAGTCCACCTGTGCGACTCTG 1430  
Db 181 CTGTCTCAGAGGTTGGACAGAAAAGAGAGAGCCCTCAGTCCACCTGTGCGACTCTG 240  
Qy 1431 CTACCTCAGAGCAGCAGAGCAAAAGCCCGTGCATCCCGCAGAGTGCCTGAGCTGCCCCA 1490  
Db 241 CTACCTCAGAGCAGCAGAGCAAAAGCCCGTGCATCCCGCAGAGTGCCTGAGCTGCCCCA 300  
Qy 1491 GCGTCAGAGCAGCAGAGCAAAAGCCCGTGCATCCCGCAGAGTGCCTGAGCTGCCCCA 1550  
Db 301 GCGTCAGAGCAGCAGAGCAAAAGCCCGTGCATCCCGCAGAGTGCCTGAGCTGCCCCA 360  
Qy 1551 TGTCCCGACAGAGCTGGAGACAGCAATAAGCTCAAGCGTTCTCTCTGTGATATCA 1610  
Db 361 TGTCCCGACAGAGCTGGAGACAGCAATAAGCTCAAGCGTTCTCTCTGTGATATCA 420  
Qy 1611 AATCAGTTTCATATTTCAGCCAGTGGAGAGATCTTACATGGCTTCTCTCTCATCAGAAG 1670  
Db 421 AATCAGTTTCATATTTCAGCCAGTGGAGAGATCTTACATGGCTTCTCTCTCATCAGAAG 480  
Qy 1671 ATGCTTTGGAATCTACAACTCTGATCTCTGATGGAGAGCAAGTATGCCAGT 1730  
Db 481 ATGCTTTGGAATCTACAACTCTGATCTCTGATGGAGAGCAAGTATGCCAGT 540  
Qy 1731 TCTCCCTGTTCAGGAAGTATCGAGAGCACTCCCGAAACAGTCTCTGATAGAGAGAG 1790  
Db 541 TCTCCCTGTTCAGGAAGTATCGAGAGCACTCCCGAAACAGTCTCTGATAGAGAGAG 600  
Qy 1791 CCAGCATCCCCAGAGAGCTGAGAGAGAGTCCCGAAACAGTCTCTGATAGAGAGAG 1850  
Db 601 CCAGCATCCCCAGAGAGTCCAGAGAGAGTCCCGAAACAGTCTCTGATAGAGAGAG 660  
Qy 1851 ATTCCGTCAGACAGCAGTGGAGAGAGTCCCGAAACAGTCTCTGATAGAGAGAG 1910  
Db 661 ATTCCGTCAGACAGCAGTGGAGAGAGTCCCGAAACAGTCTCTGATAGAGAGAG 720  
Qy 1911 GAAATGGAGAGCTGGAGAGCAATTAACACAGAGTTCCTTTTCGGCCCTTTCCACAGCC 1970  
Db 721 GAAATGGAGAGCTGGAGAGCAATTAACACAGAGTTCCTTTTCGGCCCTTTCCACAGCC 780  
Qy 1971 AGCAGACCTCAGAGTCTGTGGCC--TGGCCCTTAAGGCTGGCACTCGATATCTT 2028  
Db 781 AGCAGACCTCAGAGTCTGTGGCCCTTAAAGGCTGNNCACTCGATAT 840

RESULT 7  
BO951695 974 bp mRNA linear EST 21-AUG-2002  
LOCUS AGENCOURT\_8947130 NIH\_MGC\_101 Homo sapiens cdna clone IMAGE:6461668  
DEFINITION 5', mRNA sequence.  
ACCESSION BO951695  
VERSION BO951695.1 GI:22367173  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 974)  
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-i@mail.nih.gov](mailto:cgapbs-i@mail.nih.gov)

Tissue Procurement: ATCC  
cdna Library Preparation: Rubin Laboratory  
cdna Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LUCM2650 row: f column: 05  
High quality sequence stop: 617.  
Location/Qualifiers  
1. 974  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6461668"  
/tissue\_type="epidermoid carcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_101"  
/note="Organ: lung; Vector: pORF7; Site: 1: EcoRI; Site 2: XhoI; cdna made by oligo-dr priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."  
ORIGIN  
Query Match 23.6%; Score 786.2; DB 13; Length 974;  
Best Local Similarity 94.5%; Pred. No. 6.1e-202;  
Matches 882; Conservative 0; Mismatches 38; Indels 13; Gaps 6;  
Qy 1191 TGTCTTTAGATGAAGCTTACAGATTGTGGAAGAAAAAGACCTACTATATCTCCAACT 1250  
Db 1 TGTCTTTAGATGAAGCTTACAGATTGTGGAAGAAAAAGACCTACTATATCTCCAACT 60  
Qy 1251 TCAATTTCTGGCCAACTCTGGACTATGAGAGAGATTAAAGACAGACTGGAGCAT 1310  
Db 61 TCAATTTCTGGCCAACTCTGGACTATGAGAGAGATTAAAGACAGACTGGAGCAT 120  
Qy 1311 CAGGGCCAAAGAGCAAACTCAAGCTGCTGCACTCTGGAGAACCAAAATGAACCTGTCCCTG 1370  
Db 121 CAGGGCCAAAGAGCAAACTCAAGCTGCTGCACTCTGGAGAACCAAAATGAACCTGTCCCTG 180  
Qy 1371 CTGTCTCAGAGGTTGGACAGAAAAGAGAGAGCCCTCAGTCCACCTGTGCGACTCTG 1430  
Db 181 CTGTCTCAGAGGTTGGACAGAAAAGAGAGAGCCCTCAGTCCACCTGTGCGACTCTG 240  
Qy 1431 CTACCTCAGAGCAGCAGAGCAAAAGCCCGTGCATCCCGCAGAGTGCCTGAGCTGCCCCA 1490  
Db 241 CTACCTCAGAGCAGCAGAGCAAAAGCCCGTGCATCCCGCAGAGTGCCTGAGCTGCCCCA 300  
Qy 1491 GCGTCAGAGCAGCAGAGCAAAAGCCCGTGCATCCCGCAGAGTGCCTGAGCTGCCCCA 1550  
Db 301 GCGTCAGAGCAGCAGAGCAAAAGCCCGTGCATCCCGCAGAGTGCCTGAGCTGCCCCA 360  
Qy 1551 TGTCCCGACAGAGCTGGAGACAGCAATAAGCTCAAGCGTTCTCTCTGTGATATCA 1610  
Db 361 TGTCCCGACAGAGCTGGAGACAGCAATAAGCTCAAGCGTTCTCTCTGTGATATCA 420  
Qy 1611 AATCAGTTTCATATTTCAGCCAGTGGAGAGATCTTACATGGCTTCTCTCTCATCAGAAG 1670  
Db 421 AATCAGTTTCATATTTCAGCCAGTGGAGAGATCTTACATGGCTTCTCTCTCATCAGAAG 480  
Qy 1671 ATGCTTTGGAATCTACAACTCTGATCTCTGATGGAGAGCAAGTATGCCAGT 1730  
Db 481 ATGCTTTGGAATCTACAACTCTGATCTCTGATGGAGAGCAAGTATGCCAGT 540  
Qy 1731 TCTCCCTGTTCAGGAAGTATCGAGAGCACTCCCGAAACAGTCTCTGATAGAGAGAG 1790  
Db 541 TCTCCCTGTTCAGGAAGTATCGAGAGCACTCCCGAAACAGTCTCTGATAGAGAGAG 600  
Qy 1791 CCAGCATCCCCAGAGAGCTGAGAGAGAGTCCCGAAACAGTCTCTGATAGAGAGAG 1850

Tissue Procurement: ATCC

cdna Library Preparation: Rubin Laboratory

cdna Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LUCM2650 row: f column: 05

High quality sequence stop: 617.

Location/Qualifiers

1. 974

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:6461668"

/tissue\_type="epidermoid carcinoma, cell line"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="NIH\_MGC\_101"

/note="Organ: lung; Vector: pORF7; Site: 1: EcoRI; Site 2:

XhoI; cdna made by oligo-dr priming. Directionally cloned

into EcoRI/XhoI sites using the following 5' adaptor:

GGACGAG(G). Library constructed by Ling Hong in the

laboratory of Gerald M. Rubin (University of California,

Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and

Superscript II RT (Life Technologies). Note: this is a

NIH\_MGC Library."

Db 601 CCAGCATCCCAAGAGCTGCAGAGCGCAGCTTCAGACAGCCAGAGCGATTC 660  
 QY 1851 ATTGGGTGAGAACAGCAGCAGTGGCCAGCGCCAGAGTCCCTTTATCTCCACTGCATC 1910  
 Db 661 ATTGGGTGAGAACAGCAGCAGTGGCCAGCGCCAGAGTCCCTTTATCTCCACTGCATC 720  
 QY 1911 GAAGTGGGAGCGTGGAGACAAATACACACAGCTT---CCCTTTGGGCTTTCCACCA 1967  
 Db 721 GAAGTGGGAGCGTGGAGACAAATACACACAGCTTCCCTTTTCGGGCTTTCCACCA 780  
 QY 1968 GCCAGCAGCA-CCTCAGCAAGTCTGCTGGCTGGG--CCTTAAGGGCTGGCACTCGGA-- 2022  
 Db 781 GCCAGCAGCAGCTTCAGAACCTCTGCTGGCTGGGCTGGGCTGGGCACTTCGAAT 840  
 QY 2023 TATCTTTGGCCCCCAGACCTCTACCCCTT--CCCTGACACAGCAGCTGGTATTTGGCACA 2080  
 Db 841 ATCTTTGGGCCCCCAGACCTCTACCCCTTTCCTCTGAAACAGAGCTGGTAATTTGGCCAC 900  
 QY 2081 GAGTCTCTCACA--CTTCTACTCTGCTTCAGCC 2110  
 Db 901 CGAAGCCCTCACACCTTCTACTTGCCTTCAGCC 933

RESULT 8  
 BX110818 798 bp mRNA linear EST 07-FEB-2003  
 LOCUS BX110818 Soares fetal liver spleen INFLS Homo sapiens cDNA clone  
 DEFINITION IMAGp998F21130 ; IMAGE:127508, mRNA sequence.

ACCESSION BX110818  
 VERSION BX110818.1 GI:27836652  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 798)  
 Ebert, L., Heil, O., Hennig, S., Neubert, P., Partsch, E., Peters, M.,  
 Radelof, U., Schneider, D. and Korn, B.  
 Human Unigeneset - RZPD3  
 Unpublished (2003)

CONTACT: Ina Rofls  
 RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH  
 Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany  
 RZPD; IMAGp998F21130.  
 RZPDLIB; I.M.A.G.E. cDNA Clone Collection;  
 Human Unigeneset - RZPD3 (RZPDLIB No.972)  
 http://www.rzpd.de/CloneCards/cgi-  
 bin/showLib.pl.cgi/response?libNo=972 Contact: Ina Rofls  
 RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH  
 Heubnerweg 6, D-14059 Berlin, Germany  
 Tel: +49 30 32639 101  
 Fax: +49 30 32639 111  
 www.rzpd.de

This clone is available royalty-free from RZPD;  
 contact RZPD (clone@rzpd.de) for further information. Seq primer:  
 M13r, Primer sequence: TTTCCACAGAAACAGCTAAGAC.

FEATURES  
 source  
 1..798  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGp998F21130 ; IMAGE:127508"  
 /sex="male"  
 /dev\_stage="20 week-post conception fetus"  
 /lab\_host="DH10B (ampicillin resistant)"  
 /clone\_lib="Soares fetal liver spleen INFLS"  
 /note="Organ: Liver and Spleen; Vector: p7T73D (Pharmacia)  
 with a modified polylinker; Site\_1: Pac I; Site\_2: Eco RI;  
 1st strand cDNA was primed with a Pac I - oligo(dT) primer  
 [5' AACGGAAGATTAATTAAGATCTTTTCTTTTCTTTT 3']  
 double-stranded cDNA was ligated to Eco RI adaptors  
 (Pharmacia), digested with Pac I and cloned into the Pac I

and Eco RI sites of the modified p7T73 vector. Library  
 went through one round of normalization. Library  
 constructed by Bento Soares and M.Fatima Bonaldo."

ORIGIN  
 Query Match 23.3%; Score 777; DB 13; Length 798;  
 Best Local Similarity 99.7%; Pred. No. 1.8e-199;  
 Matches 788; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
 QY 2275 AAGCTGCCAAATGGAATTTGGAGAGAGCATCATGTCTAGAGAACAGGTTCACGGGAAGAGCT 2334  
 Db 10 AAGCTGCCAAATGGAATTTGGAGAGAGCATCATGTCTAGAGAACAGGTTCACGGGAAGAGCT 69  
 QY 2335 GGGGAAAGTGGGAGTCACTAGCTTTTGGGAGAGCATGGAATCATTTGAGTCTCTCTG 2394  
 Db 70 GGGGAAAGTGGGAGTCACTAGCTTTTGGGAGAGCATGGAATCATTTGAGTCTCTCTG 139  
 QY 2395 AGAGAAAGACACCTTGAGCTTCTATAGACAAATTTTTTTTCTTGTTCACAAAAAATTC 2454  
 Db 130 AGAGAAAGACACCTTGAGCTTCTATAGACAAATTTTTTTTCTTGTTCACAAAAAATTC 189  
 QY 2455 CCTGTAAATCTGAAAT 2514  
 Db 190 CCTGTAAATCTGAAAT 249  
 QY 2515 GTGTAAAAGCAACAGGTGGATCAACCCAGTTGTACTCTCTTAACATCTTCATTTGAGAG 2574  
 Db 250 GTGTAAAAGCAACAGGTGGATCAACCCAGTTGTACTCTCTTAACATCTTCATTTGAGAG 309  
 QY 2575 ATCAGCTAATCTTCTCTCAACAAAAATGGAAGGCGAGATGCTAGAAATCCCCCTTGAAG 2634  
 Db 310 ATCAGCTAATCTTCTCTCAACAAAAATGGAAGGCGAGATGCTAGAAATCCCCCTTGAAG 369  
 QY 2635 GAGGAAACCAATTTTATTCAGTGAATTACACATCTCTTCTTAAAGCAAGTGTC 2694  
 Db 370 GAGGAAACCAATTTTATTCAGTGAATTACACATCTCTTCTTAAAGCAAGTGTC 429  
 QY 2695 TTTGGTGTGGAGACAAAAATCCCTACCATTTCACACGTTGTGCTACTAAGAGATCTCA 2754  
 Db 430 TTTGGTGTGGAGACAAAAATCCCTACCATTTCACACGTTGTGCTACTAAGAGATCTCA 488  
 QY 2755 AATATTAGTCTTTGTCGGACCCCTTCCATAGTACACCTTAGCCCTGAGACTGAGCCAGCT 2814  
 Db 489 AATATTAGTCTTTGTCGGACCCCTTCCATAGTACACCTTAGCCCTGAGACTGAGCCAGCT 548  
 QY 2815 TGGGGTTCAGGTAGTACACCTGTTAGGACAGAGCCTAGTGGTAAATCCAAAGAGAAAT 2874  
 Db 549 TGGGGTTCAGGTAGTACACCTGTTAGGACAGAGCCTAGTGGTAAATCCAAAGAGAAAT 608  
 QY 2875 GATCCTATCCAAAGCTGATTTCACAAACCCAGCTCAGCTGACAGCGGAGGACACGAGCA 2934  
 Db 609 GATCCTATCCAAAGCTGATTTCACAAACCCAGCTCAGCTGACAGCGGAGGACACGAGCA 668  
 QY 2935 TCACCTGCTGACGAGCACCATTAGGGGCTTGGCAAGTCTACCTTAGAGCAAAACCCAGT 2994  
 Db 669 TCACCTGCTGACGAGCACCATTAGGGGCTTGGCAAGTCTACCTTAGAGCAAAACCCAGT 728  
 QY 2995 ACCTCAGACAGGAAAGTCGGGGCTTTGACCACTACCATTATCTGTTAGCCCATTTTCTAGG 3054  
 Db 729 ACCTCAGACAGGAAAGTCGGGGCTTTGACCACTACCATTATCTGTTAGCCCATTTTCTAGG 788  
 QY 3055 CATTGTGAAT 3064  
 Db 789 CATTGGAAT 798

RESULT 9  
 BG482429  
 LOCUS BG482429  
 DEFINITION 602526595F1 NIH\_MGC\_21 Homo sapiens cDNA clone IMAGE:4650430 5',  
 mRNA sequence.  
 ACCESSION BG482429  
 VERSION BG482429.1 GI:13414708

```
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 903)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLC1431 row: m column: 23
High quality sequence stop: 806.
FEATURES
source
1..903
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4650430"
/tissue_type="choriocarcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_21"
/note="Organ: placenta; Vector: pOTB7; Site: 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

```

Query Match 23.0%; Score 766.4; DB 12; Length 903;
Best Local Similarity 94.8%; Pred. No. 1.4e-196;
Matches 838; Conservative 0; Mismatches 36; Indels 10; Gaps 4;
QY 927 AGAGCTGATGCGAGCAATGGATGGTATGTTATGTTAAATCCAGCAATACCTGTCCAA 986
DB 19 AGAGCTGATGCGAGCAATGGATGGTATGTTATGTTAAATCCAGCAATACCTGTCCAA 78
QY 987 AGCTGACTTTATCCCGAGTCTCATTTCTCGTGTGCTGTGAATGACAGCTTTTGTG 1046
DB 79 AGCTGACTTTATCCCGAGTCTCATTTCTCGTGTGCTGTGAATGACAGCTTTTGTG 138
QY 1047 AGAAATTTTGGCGTGGTGGCAAAATCAGTAGATTTCATTGAGAAAGCAAAAGCTTCCA 1106
DB 139 AGAAATTTTGGCGTGGTGGCAAAATCAGTAGATTTCATTGAGAAAGCAAAAGCTTCCA 198
QY 1107 ATGATGTGTTCTAGTGCACTGTTAGCTGGGATCTCCGCTCCGCCACCATCGTATCG 1166
DB 199 ATGATGTGTTCTAGTGCACTGTTAGCTGGGATCTCCGCTCCGCCACCATCGTATCG 258
QY 1167 CCTACATCATGAAGAGATGGAATGTCATCTTTAGATGAAGCTTACAGATTGTGAAAGAAA 1226
DB 259 CCTACATCATGAAGAGATGGAATGTCATCTTTAGATGAAGCTTACAGATTGTGAAAGAAA 318
QY 1227 AAAGACCTACTATATCTCCAACTTCAATTTCTGGGCCAACTCTTGGACTATGAGAAGA 1286
DB 319 AAAGACCTACTATATCTCCAACTTCAATTTCTGGGCCAACTCTTGGACTATGAGAAGA 378
QY 1287 AGATTAGAACACAGATGAGCATCAGGCGCAAGCAAACTCAAGTGTGTCACCTGG 1346
DB 379 AGATTAGAACACAGATGAGCATCAGGCGCAAGCAAACTCAAGTGTGTCACCTGG 438
QY 1347 AGAGCCAAATGAACTGTCCCTGCTGTCTCAGAGGTGGAAGAAAGCGAGAGCGCCCC 1406
```

```

439 AGAGCCAAATGAACTGTCCCTGCTGTCTCAGAGGTGGAAGAAAGCGAGAGCGCCC 498
1407 TCAGTCCACCTGTGTCGCACTGTCTACCTCAGAGGCGAGGACAAAGGCCCGTGCATC 1466
499 TCAGTCCACCTGTGTCGCACTGTCTACCTCAGAGGCGAGGACAAAGGCCCGTGCATC 558
1467 CGCGCAGCGTGCAGCGGTCGCCAGCGTGCAGCGTGCAGCGTGCAGCGTGCAGCGTGC 1526
559 CGCGCAGCGTGCAGCGGTCGCCAGCGTGCAGCGTGCAGCGTGCAGCGTGCAGCGTGC 618
1527 TACAGGCGCTCAGTGGGCTGCACCTGTCCGAGCAGCAGGCTGGAAAGACAGCAATAAGCTCA 1586
619 TACAGGCGCTCAGTGGGCTGCACCTGTCCGAGCAGCAGGCTGGAAAGACAGCAATAAGCTCA 678
1587 AGCTTCCTCTCTCTGGATATCAAT-CAGTTTCATATTCAGCAGCAGTGGCAGCATCC 1645
679 AGCTTCCTCTCTCTGGATATCAATCAATTCATATTCAGCAGCAGTGGCAGCATCC 738
1646 TT---ACATGGCTTCTCTCTCATCAGAAG-ATGCTTTTGGAAATCACTACAAACCTTCCACTAC 1701
739 CITACATTGGCTTCTCTCTCATCAGAAGATGCTTTGGAAATCACTACAAACCTTCCACTT 798
1702 TCTGGATGGGACCAACAAGC-----TATGCCAGTTCCTCCCTGTTTCAGGAACACTATCGAG 1756
799 ACTCTGGATTGGGACCAACAAGGTTATGCCAGTTCCTCCCTGTTTCAGGAACACTATCGAA 858
1757 CAGACTCTCCGAAACCAAGCTCTGATAAGGAGGAAGCAGCATCCC 1800
859 GAGACTCTCCGAAACCAAGCTCTGATAAGGAGGAAGCAATATCCC 902
```

891 bp mRNA linear EST 15-JUL-2002  
BQ670989  
AGENCOURT\_8035377 NIH\_MGC\_102 Homo sapiens cDNA clone IMAGE:6211589  
5', mRNA sequence.  
BQ670989  
BQ670989.1 GI:21781823  
EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE NIH-MGC http://mgi.nci.nih.gov/.  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
TITLE Unpublished (1999)  
JOURNAL  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Rubin Laboratory  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLC2374 row: f column: 06  
High quality sequence stop: 633.  
Location/Qualifiers  
1..891  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6211589"  
/tissue\_type="epidermoid carcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_102"  
/note="Organ: salivary gland; Vector: pOTB7; Site: 1: XhoI;  
Site 2: EcoRI; cDNA made by oligo-dT priming.  
Directionally cloned into EcoRI/XhoI sites using the  
following 5' adaptor: GGCACGAG(G). Library constructed  
by Ling Hong in the laboratory of Gerald M. Rubin  
(University of California, Berkeley) using ZAP-cDNA



ORIGIN

synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."

Query Match 22.2%; Score 741.2; DB 13; Length 891;  
Best Local Similarity 99.4%; Pred. No. 1e-189;  
Matches 765; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

QY 1766 GAACAGCTCTGTAAAGAGGAGCCAGCTCCCAAGAGCTGCAGAGCGGAGCCT 1825  
Db 1 GAACAGCTCTGTAAAGAGGAGCCAGCTCCCAAGAGCTGCAGAGCGGAGCCT 60

QY 1826 TCAGACAGCCAGCAAGCGATTGCTTCGCTCAGAACCCAGCAGCAGTGCACGCCAG 1885  
Db 61 TCAGACAGCCAGCAAGCGATTGCTTCGCTCAGAACCCAGCAGCAGTGCACGCCAG 120

QY 1886 AGGTCCCTTTTATCTCCACTGCATCGAAGTGGAGCGTGGAGGACAATTACCAACAGC 1945  
Db 121 AGGTCCCTTTTATCTCCACTGCATCGAAGTGGAGCGTGGAGGACAATTACCAACAGC 180

QY 1946 TTCCTTTTTCGGCTTTCCACAGCCAGCAGCAGCTCAGAGTCTGCTGGCGCTT 2005  
Db 181 TTCCTTTTTCGGCTTTCCACAGCCAGCAGCAGCTCAGAGTCTGCTGGCGCTT 240

QY 2006 AAGGCTGGCTCTGCTGATATCTTGGCCCCCAGACCTCTACCCCTTCCCTGACCCAGC 2065  
Db 241 AAGGCTGGCTCTGCTGATATCTTGGCCCCCAGACCTCTACCCCTTCCCTGACCCAGC 300

QY 2066 TGGTATTTTGCAGAGTCTCAGACTTCTACTCTGCTCAGCCTCTCAGAGGAGT 2125  
Db 301 TGGTATTTTGCAGAGTCTCAGACTTCTACTCTGCTCAGCCTCTCAGAGGAGT 360

QY 2126 GCCAGTACTCTGCTCAGAGTCTGAGCCAGCTGCGCCACTTGGGAGACCAAGTCTATTCT 2185  
Db 361 GCCAGTACTCTGCTCAGAGTCTGAGCCAGCTGCGCCACTTGGGAGACCAAGTCTATTCT 420

QY 2186 GTGCCAGCGGCGAGAGCCAGTGAAGAGTCAAGCTGCTCGCGCGAGCTGGCATGAAGAG 2245  
Db 421 GTGCCAGCGGCGAGAGCCAGTGAAGAGTCAAGCTGCTCGCGCGAGCTGGCATGAAGAG 480

QY 2246 AGCCCTTTTGAAGAGCTTTAAAGCGAGAGCTGCCAAATGGAATTTGGAGAGAGCATC 2305  
Db 481 AGCCCTTTTGAAGAGCTTTAAAGCGAGAGCTGCCAAATGGAATTTGGAGAGAGCATC 540

QY 2306 ATGTCAAGAGACAGTCAAGGAGAGCTGGGAAAGTGGGAGTCAAGTCAAGTCTTTTGG 2365  
Db 541 ATGTCAAGAGACAGTCAAGGAGAGCTGGGAAAGTGGGAGTCAAGTCAAGTCTTTTGG 600

QY 2366 GGCAGCATGGAATCATTTAGGCTCTCTGAGAGAAAGACACTTGTGACTTCTATAGACA 2425  
Db 601 GGCAGCATGGAATCATTTAGGCTCTCTGAGAGAAAGACACTTGTGACTTCTATAGACC 660

QY 2426 ATTTTCTTTTCTGTTCACAAAAAATTCCTGTAAATCTGAAATATATATATGTACATA 2485  
Db 661 ATTTTCTTTTCTGTTCACAAAAAATTCCTGTAAATCTGAAATATATATATGTACATA 720

QY 2486 CATATATATTTT-GGAAATGGAGCTAT-GGTGTAAAGCAACAGGTGG 2533  
Db 721 CATATATATTTTTCGGAAATGGAGCTATGGGGGTAAAGCAACAGGTGG 770

RESULT 11  
LOCUS BM906608  
DEFINITION AGENCOURT\_6621858 NIH\_MGC\_125 Homo sapiens cDNA clone IMAGE:5722534  
5', mRNA sequence.  
ACCESSION BM906608  
VERSION BM906608.1 GI:19356987  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1148)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
TITLE Unpublished (1999)  
JOURNAL  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: Invitrogen  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
Plate: L1AM12708 Row: 1 Column: 23  
High quality sequence start: 8  
High quality sequence stop: 661.  
Location/Qualifiers  
1. 1148  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5722534"  
/lab\_host="DH10B"  
/clone\_lib="NIH\_MGC\_125"  
/note="Organ: ovary (pool of 3); Vector: pCMV-SPORT6;  
Site 1: EcoRV (destroyed); Site 2: NotI; RNA source pool of three ovaries, from females ranging in age from 38 to 49 yo. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 2.1 kb, insert size range 1-3.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 036."

ORIGIN

Query Match 22.1%; Score 736.4; DB 12; Length 1148;  
Best Local Similarity 97.8%; Pred. No. 2.4e-188;  
Matches 757; Conservative 0; Mismatches 16; Indels 1; Gaps 1;

QY 164 TCTCAGGGAATTTGGAGTCTGGCGGCCCAAAAGCTTTTCAGTCCAGTGTAAAGCTTTGG 223  
Db 3 TCTCAGGGAATTTGGAGTCTGGCGGCCCAAAAGCTTTTCAGTCCAGTGTAAAGCTTTGG 62

QY 224 AGCGCGGAGCAAAAGGTAAAGATGATGTAATCGCTGGCTGTCCAAAGCATCTTTTGT 283  
Db 63 AGCGCGGAGCAAAAGGTAAAGATGATGTAATCGCTGGCTGTCCAAAGCATCTTTTGT 122

QY 284 TGTGGAATGTTATTCAGTCATCTTTATGATCAATGTAATGTGAGGGGTGCTTTGTGA 343  
Db 123 TGTGGAATGTTATTCAGTCATCTTTATGATCAATGTAATGTGAGGGGTGCTTTGTGA 182

QY 344 CGGAGTCTTTTGAAGAGCAGCATCAAGCGGAAAGAGAGACATTCATCTGGAGGGCT 403  
Db 183 CGGAGTCTTTTGAAGAGCAGCATCAAGCGGAAAGAGAGACATTCATCTGGAGGGCT 242

QY 404 CTTCCTGAAATGGGTTTAACTCTCTTTTCCAGTCCACCCAGCTGACCTCATACAC 463  
Db 243 CTTCCTGAAATGGGTTTAACTCTCTTTTCCAGTCCACCCAGCTGACCTCATACAC 302

QY 464 TTTTGTAGTCAATGGAGTGGCTGAGCCTTTGAGCACACCACTTACATCATCTGGCAAA 523  
Db 303 TTTTGTAGTCAATGGAGTGGCTGAGCCTTTGAGCACACCACTTACATCATCTGGCAAA 362

QY 524 TTAAGAGAGAGTGGGAAAGAGGACTTATTTGTTCTATGGCCCATGAGATGATTGGA 583  
Db 363 TTAAGAGAGAGTGGGAAAGAGGACTTATTTGTTCTATGGCCCATGAGATGATTGGA 422

QY 584 CTCAATTTGTTTACTGAGAGTGTGCTGCTCTGCTGAAAGTGGAAAGGAAAGTCTCTGC 643  
Db 423 CTCAATTTGTTTACTGAGAGTGTGCTGCTCTGCTGAAAGTGGAAAGGAAAGTCTCTGC 482

QY 644 TAATGTATAGCCGCCCATTTTGTGGAATACATATACATCCACATTTTGGAGCCATTA 703

```
Db 483 TAATTGATAGCGGCAATTGTGGATACAAATACATCCACATTTTGGAGCCATTATA 542
Qy 704 TCAACTGCTCCAGACTTATGAAGCGAAGGTTCCACAGCAAAAGTGTAAATACAGAGC 763
Db 543 TCAACTGCTCCAGACTTATGAAGCGAAGGTTCCACAGCAAAAGTGTAAATACAGAGC 602
Qy 764 TCAATCAGCAATTCAGCGAAACATAAGGTTGACATTTGATTCAGAGTCAGAGGTTGAGTTT 823
Db 603 TCAATCAGCAATTCAGCGAAACATAAGGTTGACATTTGATTCAGAGTCAGAGGTTGAGTTT 662
Qy 824 AGATCAAAAGTCCCAAGATGTTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 882
Db 663 AGATCAAAAGTCCCAAGATGTTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 722
Qy 883 CTGGGTAAACTCGAGAGAGCTTCAACTCTGTTCACTCTGTTGCTGAGGAGCTGAT 936
Db 723 CTGGGAAAACCTGGAAGAAGCTTCACTTTGGCCACCCGCTTGCCGGTGGGCAT 776

RESULT 12
AW963248
LOCUS 740 bp mRNA linear EST 01-JUN-2000
DEFINITION EST375321 MAGE resequences, MAGH Homo sapiens cDNA, mRNA sequence.
ACCESSION AW963248
VERSION AW963248.1 GI:8153084
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Hegde,P., Qi,R., Abernathy,K., Dharap,S., Gaspard,R., Gay,C.,
Holte,I.B., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and
Quackenbush,J.
TITLE Assessment of gene expression patterns in a model of colon tumor
metastasis using a 19,200 element cDNA microarray
JOURNAL Unpublished (2000)
COMMENT Contact: John Quackenbush
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 3528
Fax: 301 838 0208
Email: johnqu@igr.org
Plate: 185
Seq primer: Reverse.
FEATURES
source
1..740
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="MAGE resequences, MAGH"
/notes="Vector: pBluescriptSKm"

ORIGIN
Query Match 21.5%; Score 717.8; DB 10; Length 740;
Best Local Similarity 98.9%; Pred. No. 2.2e-183;
Matches 733; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

Qy 2511 TATGGTGTAAAGCAACAGGTGGATCAACCCAGTTGTTACTCTCTTAACATCTGCTTTG 2570
Db 1 TATGGTGTAAAGCAACAGGTGGATCAACCCAGTTGTTACTCTCTTAACATCTGCTTTG 60
Qy 2571 AGAGATCAGCTAATCTCTCTCAACAAAATGGAAGGCGCAGATGCTAGATCCCCCTTA 2630
Db 61 AGAGATCAGCTAATCTCTCTCAACAAAATGGAAGGCGCAGATGCTAGATCCCCCTTA 120
Qy 2631 GACGGAGGAAACCAATTTTATTCAGTGAATACATCTCTCTGTTCTTAAAAAGCAAG 2690
Db 121 GACGGAGGAAACCAATTTTATTCAGTGAATACATCTCTCTGTTCTTAAAAAGCAAG 180
Qy 2691 TGTCTTTGTTGGAGGACAAAATCCCTACATTTTCCAGTTGTTGCTACTAAGAGAT 2750
Db 181 TGTCTTTGTTGGAGGACAAAATCCCTACATTTT-CAGTTGTGCTACTAAGAGAT 239
```

```
Qy 2751 CTCAAATATTAGTCTTTGTCCGGACCCCTTCCATAGTACACCTTAGCGCTGAGACTGAGCC 2810
Db 240 CTCAAATATTAGTCTTTGTCCGGACCCCTTCCATAGTACACCTTAGCGCTGAGACTGAGCC 299
Qy 2811 AGCTTGGGGGTCAGGTAGTAGAGCCCTGTTAGGACAGAGCCTAGTGTAAATCCCAAGAG 2870
Db 300 AGCTTGGGGGTCAGGTAGTAGAGCCCTGTTAGGACAGAGCCTAGTGTAAATCCCAAGAG 359
Qy 2871 AAATGATCTCTATCCAAAGCTGATTCACAAACCCACGCTCACCTGACAGCCGAGGACAG 2930
Db 360 AAATGATCTCTATCCAAAGCTGATTCACAAACCCACGCTCACCTGACAGCCGAGGACAG 419
Qy 2931 AGCATCACTCTGCTGGACGAGCACTTATAGGGGCTTGCACCAAGGCTTACCTTAGAGCAAAAC 2990
Db 420 AGCATCACTCTGCTGGACGAGCACTTATAGGGGCTTGCACCAAGGCTTACCTTAGAGCAAAAC 479
Qy 2991 CAGTACCTCAGACAGGAAAGTCGGGGCTTTGACACATACCATATCTGTTAGCCCATTTTC 3050
Db 480 CAGTACCTCAGACAGGAAAGTCGGGGCTTTGACACATACCATATCTGTTAGCCCATTTTC 539
Qy 3051 TAGGCATTTGCAATAGGTAGGTAGTACCTAGTACACCTTTTTCAGACCAATTCAAACTGCTAT 3110
Db 540 TAGGCATTTGCAATAGGTAGGTAGTACCTAGTACACCTTTTTCAGACCAATTCAAACTGCTAT 599
Qy 3111 GCACAAAATTCCTGGGCTTAGATGAGATTAATTTTTTTTCTCTCAGCTTTATGAAG 3170
Db 600 GCACAAAATTCCTGGGCTTAGATGAGATTAATTTTTTTTCTCTCAGCTTTATGAAG 659
Qy 3171 AGAAGGGAACTGCTAGGATTCAGCTGAACCCAGGACCTGGCAACATCAGATTTA 3230
Db 660 AGAAGGGAACTGCTAGGATTCAGCTGAACCCAGGACCTGGCAACATCAGATTTA 719
Qy 3231 AGCTAAGGTTGGAGGCTTAAC 3251
Db 720 AGCTAAGGTTGGAGGCTTAAC 740

RESULT 13
CA944683/c
LOCUS 720 bp mRNA linear EST 30-DEC-2002
DEFINITION UI-CF-FN0-af-r-i-08-0-UI.s1 UI-CF-FN0 Homo sapiens cDNA clone
UI-CF-FN0-af-r-i-08-0-UI 3', mRNA sequence.
ACCESSION CA944683
VERSION CA944683.1 GI:27433163
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE Ronaldo,M.F., Lennon,G. and Soares,M.B.
AUTHORS Normalization and subtraction: two approaches to facilitate gene
TITLE discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PUBMED 8889548
COMMENT Contact: McCray, PB
McCray Lab
University of Iowa
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
Tel: 319 356 4866
Fax: 319 356 7171
Email: paul-mccray@uiowa.edu
Tissue Procurement: Dr. M. J. Welsh, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com) or from Open Biosystems
(www.openbiosystems.com).
Seq primer: M13 FORWARD
POLYA=Yes.
```

```
FEATURES
source
Location/Qualifiers
1..720
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="UI-CF-FNO-af-i-08-0-UI"
/tissue_type="Human Lung Epithelial cells"
/lab_host="DH10B (Life Technologies) (TI phage resistant)"
/clone_lib="UI-CF-FNO"
/notes="Organ: Lung; Vector: pMT73-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
UI-CF-FNO is a subcloned cDNA library derived from two
normalized Human lung epithelial cell libraries (EN1 and
DUI1). The library was subcloned according to
Ronald, Lennon and Soares, Genome Research, 6:791-806,
1996. For additional information, contact:
bento-soares@uiowa.edu
TAG TISSUE=Lung Epithelial Cells Tissue nos 359-368
TAG LIB=UI-CF-FNO
TAG_SEQ=GGCTGTAGGC"

ORIGIN
Query Match 20.7%; Score 688.4; DB 14; Length 720;
Best Local Similarity 99.6%; Pred. No. 2.2e-175;
Matches 700; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 2630 AGACGGAGAAACCATTTTATTCAGTGAATTACATCTCTGTCTTAAAAAGCAA 2689
DB 720 AGACGGAGAAACCATTTTATTCAGTGAATTACATCTCTGTCTTAAAAAGCAA 661
QY 2690 GTGCTTTTGGTGTGGAGCAAAATCCCTACCATTTCACGTTGTGCTACTAAGAGA 2749
DB 660 GTGCTTTTGGTGTGGAGCAAAATCCCTACCATTTCACGTTGTGCTACTAAGAGA 602
QY 2750 TCTCAATATAGTCTTGTGCGGACCTTCCATAGTACACCTTAGCGCTGAGACTGACC 2809
DB 601 TCTCAATATAGTCTTGTGCGGACCTTCCATAGTACACCTTAGCGCTGAGACTGACC 542
QY 2810 CAGCTTGGGGGTGAGTAGTACCTGTTAGGACAGAGCCTGAGTGGTAAATCAAGA 2869
DB 541 CAGCTTGGGGGTGAGTAGTACCTGTTAGGACAGAGCCTGAGTGGTAAATCAAGA 482
QY 2870 GAAATGATCTTATCAAAAGCTGATTCACAAACCCACGCTACCTGACAGCCGAGGACAC 2929
DB 481 GAAATGATCTTATCAAAAGCTGATTCACAAACCCACGCTACCTGACAGCCGAGGACAC 422
QY 2930 GAGCATCACTCTGCGAGCGACCAATTAGGGGCTTCCAGGCTACCTTAGAGCAAC 2989
DB 421 GAGCATCACTCTGCGAGCGACCAATTAGGGGCTTCCAGGCTACCTTAGAGCAAC 362
QY 2990 CCAGTACCTCAGACAGGAAAGTTCGGGGCTTGGACCACTACCATATCTGGTAGCCCATTTT 3049
DB 361 CCAGTACCTCAGACAGGAAAGTTCGGGGCTTGGACCACTACCATATCTGGTAGCCCATTTT 302
QY 3050 CTAGGCATTTGAATAGTAGTAGTCTAGTACACATTTTCCAGCAATTCAAACTGTCTA 3109
DB 301 CTAGGCATTTGAATAGTAGTAGTCTAGTACACATTTTCCAGCAATTCAAACTGTCTA 242
QY 3110 TGCACAAATTCCTGGGCTAGATGAGATTAATTTTTTTCTCTCAGCTTTATGAA 3169
DB 241 TGCACAAATTCCTGGGCTAGATGAGATTAATTTTTTTCTCTCAGCTTTATGAA 182
QY 3170 GAGAGGGAACACTGTCTAGGATTCAGCTGAACACCAAGAACCTTGGCAAATCACGATTT 3229
DB 181 GAGAGGGAACACTGTCTAGGATTCAGCTGAACACCAAGAACCTTGGCAAATCACGATTT 122
QY 3230 AAGCTTAAGGTTGGGAGGCTACGAGTCTACCTCCCTCTTTGTAATCAAGAAATTTGTTA 3289
DB 121 AAGCTTAAGGTTGGGAGGCTACGAGTCTACCTCCCTCTTTGTAATCAAGAAATTTGTTA 62
QY 3290 AAATGGGATTTGCAATCTCTTAAATAAGATGAACCTTGGTTTC 3332
DB 61 AAATGGGATTTGCAATCTCTTAAATAAGATGAACCTTGGTTTC 19
```

```
RESULT 14
B8897795
LOCUS
DEFINITION
60143845JF1 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:3923247 5',
mRNA sequence.
B8897795
VERSION
B8897795.1 GI:10363618
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 920)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
AUTHORS
Contact: Robert Strausberg, Ph.D.
Email: cga@psh-remail.nih.gov
Tissue Procurement: ATCC/DCTD/DTF
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM9759 row: b column: 16
High quality sequence stop: 678.
FEATURES
source
1..920
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3923247"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 72"
/notes="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2 kb. Library constructed by Life
Technologies."
```

```
ORIGIN
Query Match 19.2%; Score 641.2; DB 10; Length 920;
Best Local Similarity 95.0%; Pred. No. 1.7e-162;
Matches 709; Conservative 0; Mismatches 23; Indels 14; Gaps 4;

QY 1096 AAAAGCCTCCAAATGGATGTGTTCTAGTGCACTGTTAGCTGGGATCTCCGCTCCGCCAC 1155
DB 5 AAAGCCTCCGGTGGATGTGTTCTAGTGCACTGTTAGCTGGGATCTCCGCTCCGCCAC 64
QY 1156 CATCGCTATCGCTTACATCATGAAGAGATGGACATGCTTTTAGATGAAGCTTACAGATT 1215
DB 65 CATCGCTATCGCTTACATCATGAAGAGATGGACATGCTTTTAGATGAAGCTTACAGATT 124
QY 1216 TGTGAAAGAAAAAGACCTACTATATCTCCAAACTCAATTTTCTGGGCAACTCTCTGGA 1275
DB 125 TGTGAAAGAAAAAGACCTACTATATCTCCAAACTCAATTTTCTGGGCAACTCTCTGGA 184
QY 1276 CTATGAGAGAGATTAAAGACCACTGAGCATCAGGCCCAAGAGCAAACTCAAGCT 1335
DB 185 CTATGAGAGAGATTAAAGACCACTGAGCATCAGGCCCAAGAGCAAACTCAAGCT 244
QY 1336 GCTGCACCTGAGAGAGCAAAATGAACCTGTCCCTGTCTCTCAGAGGGTGGACAGAAAAG 1395
DB 245 GCTGCACCTGAGAGAGCAAAATGAACCTGTCCCTGTCTCTCAGAGGGTGGACAGAAAAG 304
QY 1396 CGAGAGCCCTCTAGTCCACCTGTGTGCCGACTCTGTACTCTAGAGGAGCAGACAGAG 1455
DB 305 CGAGAGCCCTCTAGTCCACCTGTGTGCCGACTCTGTACTCTAGAGGAGCAGACAGAG 364
QY 1456 GCCCGTGCATCCCGCCAGCGTGCAGCGTGCAGCGTGCAGCGTGCAGCGTGCAGCGTGCAGCG 1515
```

Db 365 GCCCGTGCATCCCG-----CCAGCGTGCCAGCGTGCAGCGTGCCTGTTAGAGGA 415  
Qy 1516 CAGCCCGCTGGTACAGCGCTCAGTGGGTGCACCTGTCCGAGACAGCGCTGGAGACAG 1575  
Db 416 CAGCCCGCTGGTACAGCGCTCAGTGGGTGCACCTGTCCGAGACAGCGCTGGAGACAG 475  
Qy 1576 CAATAGCTCAAGCGTTCCTCTCTCTGATATCAATCAGTTTCATATTCAGCCAGCAT 1635  
Db 476 CAATAGCTCAAGCGTTCCTCTCTCTGATATCAATCAGTTTCATATTCAGCCAGCAT 535  
Qy 1636 GGCAGCATCCTTACATGGCTTCTCCTCATCAGAGATGCTTTGGAACTACTACAAACCTTC 1695  
Db 536 GGCAGCATCCTTACATGGCTTCTCCTCATCAGAGATGCTTTGGAACTACTACAAACCTTC 595  
Qy 1696 CACTACTCTGGATGGGACCAAGACTATGCCAGTTCCTCCCTGTTCAGGAACCTATCGGA 1755  
Db 596 CACTACTCTGGATGGGACCAAGACTATGCCAGTTCCTCCCTGTTCAGGAACCTATCGGA 654  
Qy 1756 GCAGACTCCCGAAACCAAGTCTCTGATAGGAGGAAGCCAGCATCCCAAGAGCTGCAGAC 1815  
Db 655 GCAGACTCCCGAACCC---AGTCTGATAGGAGGAAGCCAGCATCCCAAGAGGTCGACAC 711  
Qy 1816 GGCAGGCTTCAGAGCCAGCAGCA 1841  
Db 712 CG-CAGGCTTCAGGAAGCCAGAGCA 736

RESULT 15  
BX479029 663 bp mRNA linear EST 04-SEP-2003  
LOCUS DXFZp686J22208\_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone  
DEFINITION DXFZp686J22208 5', mRNA sequence.  
ACCESSION BX479029  
VERSION BX479029.1 GI:31914415  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 663)  
AUTHORS Bahr A., Lauber J., Mewes H.W., Weil B., Amid C., Oesanger A.,  
Pobo G., Han M. and Wiemann S.  
EST (Bahr A., Lauber J., Mewes H.W., Weil B., et al.)  
JOURNAL Unpublished (2003)  
COMMENT Contact: MIPS

MIPS  
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany  
This is the 5' sequence of the clone insert  
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing  
consortium of the German Genome Project.  
NO sl sequence available.  
This clone (DKFZp686J22208) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES  
source  
1. .663  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="DKFZp686J22208"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/clone\_lib="686 (synonym: hlcc3)"  
/note="vector: pTriplex2; Site\_1: SfIIA; Site\_2: SfiIB;  
cDNA-collection"

ORIGIN  
Query Match 19.0%; Score 632; DB 13; Length 663;  
Best Local Similarity 98.6%; Pred.No. 4.7e-160;  
Matches 647; Conservative 0; Mismatches 7; Indels 2; Gaps 1;

Qy 1062 GGTGGACAAATCAGTAGATTTCATT--GAGAAAGCAAAGCCTCCATGATGCTGTCT 1119  
Db 2 GGTGGACAAATCAGTAGATTTCANTTGACGAAAGCAAAGCCTCCATGATGCTGTCT 61  
Qy 1120 AGTGCACATGTTTAGTGTGGATCTCCGCTCCGACCACTCGCTATCGCTACATCATGAA 1179  
Db 62 AGTGCACATGTTTAGTGTGGATCTCCGCTCCGACCACTCGCTATCGCTACATCATGAA 121  
Qy 1180 GAGGATGGACATGCTTTAGATGAAGCTTACAGATTGTGAAAGAAAAGACCTACTAT 1239  
Db 122 GAGGATGGACATGCTTTAGATGAAGCTTACAGATTGTGAAAGAAAAGACCTACTAT 181  
Qy 1240 ATCTCCAAACTTCAATTTCTGGGCCAACTCTCGACTATCAGAAAGAGATTAAAGAACCA 1299  
Db 182 ATCTCCAAACTTCAATTTCTGGGCCAACTCTCGACTATCAGAAAGAGATTAAAGAACCA 241  
Qy 1300 GACTGGAGCATCAGGGCCAAAGAGCAAACTCAAGCTGTGCACTGGAGAAAGCCAAATGA 1359  
Db 242 GACTGGAGCATCAGGGCCAAAGAGCAAACTCAAGCTGTGCACTGGAGAAAGCCAAATGA 301  
Qy 1360 ACTGTTCCTGTCTGTCTCAGAGGTGGACAGAAAGCGAGACGCCCTCAGTCCACCTG 1419  
Db 302 ACTGTTCCTGTCTGTCTCAGAGGTGGACAGAAAGCGAGACGCCCTCAGTCCACCTG 361  
Qy 1420 TGCCGACTCTGCTACTCTCAGAGGCAGAGCAAAAGGCCGCTGATCCGCCAGCGTGCC 1479  
Db 362 TGCCGACTCTGCTACTCTCAGAGGCAGAGCAAAAGGCCGCTGATCCGCCAGCGTGCC 421  
Qy 1480 CAGCGTGCCAGCGTGCAGCCGCTGCTGTTAGAGACAGCCCGCTGGTACAGGCGCTCAG 1539  
Db 422 CAGCGTGCCAGCGTGCAGCCGCTGCTGTTAGAGACAGCCCGCTGGTACAGGCGCTCAG 481  
Qy 1540 TGGGCTGCACCTGTCCGAGACAGCGCTGGAAGACAGCAATAAGCTCAAGCGTTCCTTCTC 1599  
Db 482 TGGGCTGCACCTGTCCGAGACAGCGCTGGAAGACAGCAATAAGCTCAAGCGTTCCTTCTC 541  
Qy 1600 TCTGGATATCAATCAGTTTCATATTCAGCCAGCATGCGCAGCATCTTACATGCTTCTC 1659  
Db 542 TCTGGATATCAATCAGTTTCATATTCAGCCAGCATGCGCAGCATCTTACATGCTTCTC 601  
Qy 1660 CTCATCAGAAAGATGCTTTGGAACTACTACAAACCTTCCACTACTCTGGATGGAGCA 1715  
Db 602 CTCATCAGAAAGATGCTTTGGAACTACTACAAACCTTCCACTACTCTGGATGGAGCA 657

Search completed: February 28, 2004, 00:53:58  
Job time : 8130 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 27, 2004, 22:38:54 ; Search time 6046 Seconds

(without alignments)

3706.310 Million cell updates/sec

Title: US-09-964-277-21

Perfect score: 517

Sequence: 1 MLPLSLQTVFLYFWNRR.....LGKVGSGFSGSMELIEVS 517

Scoring table:

OLIGO  
Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Word size: 1

Total number of hits satisfying chosen parameters: 6934743

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:  
-O=/cpn2.1/USPTO.spool/US09964277/runat.25022004.085401.2209/app.query.fasta.1.711  
-DB=GenEmbl -OEXT=fastcap -SUFFIX=olip2n.rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45  
-DOCAALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptb  
-NORM=ext -HEADSIZE=500 -MINLEN=50 -MAXLEN=2000000000  
-USER=US09964277 @CGN.1.1.3635 @runat.25022004.085401.2209 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=60 -DELEXT=7

Database :

GenEmbl:  
1: gb.ba:\*  
2: gb.htg:\*  
3: gb.in:\*  
4: gb.om:\*  
5: gb.ov:\*  
6: gb.pat:\*  
7: gb.ph:\*  
8: gb.pl:\*  
9: gb.pr:\*  
10: gb.ro:\*  
11: gb.sts:\*  
12: gb.sy:\*  
13: gb.un:\*  
14: gb.vi:\*  
15: em.ba:\*  
16: em.fun:\*  
17: em.hum:\*  
18: em.in:\*  
19: em.mu:\*  
20: em.om:\*  
21: em.or:\*  
22: em.ov:\*  
23: em.pat:\*  
24: em.ph:\*  
25: em.pl:\*  
26: em.ro:\*  
27: em.sts:\*  
28: em.un:\*

RESULT 1

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	517	100.0	1935	9	AY038927 Homo sapi
2	517	100.0	3332	6	AX441229 Sequence
3	488	94.4	1998	6	AX260342 Sequence
4	488	94.4	2102	6	AX713989 Sequence
5	488	94.4	2102	9	AK055973 Homo sapi
6	488	94.4	2732	6	AX180875 Sequence
7	488	94.4	3059	6	AX278461 Sequence
8	488	94.4	3104	6	AX405700 Sequence
9	488	94.4	3496	6	AX441210 Sequence
10	488	94.4	3521	9	AB052156 Homo sapi
11	488	94.4	3544	6	AX260340 Sequence
12	488	94.4	3566	9	AF506796 Homo sapi
13	488	94.4	3766	6	AX374994 Sequence
14	488	94.4	4790	6	BD171157 Novel gen
15	488	94.4	4790	6	BD183422 Novel gen
16	488	94.4	4790	9	AB051487 Homo sapi
17	472	91.3	5450	6	AX482439 Sequence
18	472	91.3	5450	6	AX482478 Sequence
19	455	88.0	2118	6	AX099933 Sequence
20	408	78.9	2200	6	AX921915 Sequence
21	394	76.2	172206	9	AC007619 Homo sapi
22	394	76.2	188344	2	AC131617 Homo sapi
23	387	74.9	2807	9	BC031643 Homo sapi
24	350	67.7	5111	6	AX482372 Sequence
25	304	58.8	2071	6	AX921917 Sequence
26	304	58.8	3284	9	BC042101 Homo sapi
27	178	34.4	1916	6	AX835305 Sequence
28	178	34.4	1916	9	AK098310 Homo sapi
29	66	12.8	242590	2	AC133722 Rattus no
30	66	12.8	244605	2	AC097818 Rattus no
31	66	12.8	250782	2	AC128093 Rattus no
32	63	12.2	2756	6	AX482444 Sequence
33	63	12.2	3861	10	AF345954 Mus muscu
34	63	12.2	4026	10	AF345953 Mus muscu
35	63	12.2	4827	10	AF345952 Mus muscu
36	63	12.2	4874	10	AB052157 Mus muscu
37	63	12.2	4943	10	BC059232 Mus muscu
38	63	12.2	4975	10	BC057321 Mus muscu
39	63	12.2	4992	10	AF345951 Mus muscu
40	61	11.8	201474	2	AC126692 Mus muscu
41	61	11.8	236589	2	AC118035 Mus muscu
42	50	9.7	301	6	AX524414 Sequence
43	50	9.7	301	6	AX553152 Sequence
44	45	8.7	749	6	AX868469 Sequence
45	45	8.7	749	6	BD148531 Primer fo

ALIGNMENTS

AY038927 1935 bp mRNA linear PRI 01-AUG-2001  
LOCUS Homo sapiens truncated MAPK phosphatase 7 (MKP7) mRNA, complete  
DEFINITION cds, alternatively spliced.  
ACCESSION AY038927  
VERSION AY038927.1 GI:15072488  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 1935)  
AUTHORS Montpetit, A., Boily, G. and Sinnett, D.  
TITLE A detailed transcriptional map of the chromosome 12p12 tumor  
suppressor locus  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1935)  
AUTHORS Montpetit, A., Boily, G. and Sinnett, D.  
TITLE Direct Submission  
JOURNAL Submitted (07-JUN-2001) Hemato-Oncology, Hopital Ste-Justine, 3175  
Cote-Ste-Catherine, Montreal, QC H3T 1C5, Canada  
FEATURES  
source  
1..1935  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
gene 1..1935  
/gene="MKP7"  
CDS 28..459  
/gene="MKP7"  
/note="minor splice variant"  
/codon\_start=1  
/product="truncated MAPK phosphatase 7"  
/protein\_id="AAK69770.1"  
/db\_xref="GI:15072489"  
/translation="MAHEMIGTQIVTERLVALLBSGTEKVLIDSRPFVBNVSHILE  
ALINCSKLMKRLQDDKVLITELIQHSAKHKVDIDCSQKVVDYDQSSQDVASLSSDC  
FLTVLLGKLEKFSVHLLAGADAENWDLCKVCKOQYLSKA"  
ORIGIN  
Alignment Scores:  
Pred. No.: 0 Length: 1935  
Score: 517.00 Matches: 517  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0  
US-09-964-277-21 (1-517) x AY038927 (1-1935)  
QY 1 MetLeuProLeuSerLeuGlnThrValPheSerLeuTyrPheTyrValAsnTyrArgArg 20  
DB 308 ATGTTGCTCTCTCTTCAGAGCTGTTTCTCACTGTAATCTTGGGTAAACTGGAGAAGA 367  
QY 21 AlaSerThrLeuPheThrCysLeuGlnGluLeuMetGlnAsnGlyIleGlyTyrVal 40  
DB 368 GCTTCAACTCTGTTACCTGCTGCGAGGAGCTGATGACAGAAATGGGATGGTTATG 427  
QY 41 LeuAsnAlaSerAsnThrCysProLysProAspPheIleProGluSerHisPheLeuArg 60  
DB 428 TTAATGCCAGCAATACCTGTCGCAAGCTGACTTTATCTCCCGAGTCTCATTTCTCGGT 487  
QY 61 ValProValAsnAspSerPheCysGluIleLeuProTyrTyrLeuAspLysSerValAsp 80  
DB 488 GTGCTGTGAATGACAGCTTTTGTGAGAAATTTTGGCGTGGTTGGACAAATCAGTAGAT 547  
QY 81 PheIleGluLysAlaLysAlaSerAsnGlyCysValLeuValHisCysLeuAlaGlyIle 100  
DB 548 TTCATTGAGAAAGCAAAAGCTCCCAATGATGTGTTCTAGTCACTGTTTAGCTGGGATC 607  
QY 101 SerArgSerAlaThrIleAlaIleAlaTyrIleMetLysArgMetAspMetSerLeuAsp 120  
DB 608 TCCGCTCCGCCCACTCGGTATCGCTACATCATGAGAGGATGGACATGCTCTTAGAT 667

QY 121 GluAlaTyrArgPheValLysGluLysArgProThrIleSerProAsnPheAsnPheLeu 140  
DB 668 GAAGCTTACAGATTGTGAAAGAAAAAGACCTACTATATCTCCAAACTTCAATTTCTG 727  
QY 141 GlyGlnLeuLeuAspTyrGluLysIleLysAsnGlnThrGlyAlaSerGlyProLys 160  
DB 728 GGCCTACTCTGGACTATGAGAGAGATTAGAACCCAGACTGGAGCATCAGGCCAAAG 787  
QY 161 SerLysLeuLysLeuLeuHisLeuGluLysProAsnGluProValProAlaValSerGlu 180  
DB 788 AGCAAACTCAAGCTGCTGCACCTGGAGAAGCAATGAACCTGTCCCTGTCTCAGAG 847  
QY 181 GlyGlyGlnLysSerGluThrProLeuSerProProCysAlaAspSerAlaThrSerGlu 200  
DB 848 GTTGACAGAAAGAGGAGACGCCCTCAGTCCACCCTGTGCCGACTCTGTACTCTAGAG 907  
QY 201 AlaAlaGlyGlnArgProValHisProAlaSerValProSerValProSerValGlnPro 220  
DB 908 GCAGCAGGACAAAGGCCCGTGCATCCCGCAGCGTGCAGCGTGCAGCGTGCAGCGC 967  
QY 221 SerLeuLeuGluAspSerProLeuValGlnAlaLeuSerGlyLeuHisLeuSerAlaAsp 240  
DB 968 TCGCTGTGTAGAGGACAGCCCGCTGTACAGCGCTCAGTGGCGTGCACCTGTCCGAGAC 1027  
QY 241 ArgLeuGluAspSerAsnLysLeuLysArgSerPheSerLeuAspLysSerValSer 260  
DB 1028 AGCTGGAGACAGCAATAAGCTCAAGCGTTCTCTCTCTGGATATCAAAATCAGTTTCA 1087  
QY 261 TyrSerAlaSerMetAlaAlaSerLeuHisGlyPheSerSerSerGluAspAlaLeuGlu 280  
DB 1088 TATTTCAGCAGCAGTGGCAGCATCTTACATGGCTTCTCTCATCAGAGATGCTTTGGAA 1147  
QY 281 TyrTyrLysProSerThrThrLeuAspGlyThrAsnLysLeuCysGlnPheSerProVal 300  
DB 1148 TACTACAACCTTCCACTACTCTGGATGGGACCAACAGCTATCCAGTTCCTCCCTGTT 1207  
QY 301 GlnGluLeuSerGluGlnThrProGluThrSerProAspLysGluGluAlaSerIlePro 320  
DB 1208 CAGGAATATCGAGCAGACTCCGAAACCACTGCTGATAGGAGAGGAGCAGCATCCCC 1267  
QY 321 LysLysLeuGlnThrAlaArgProSerAspSerGlnSerLysArgLeuHisSerValArg 340  
DB 1268 AAGAAGCTTGACAGCCGCGGCTTTCAGACAGCAGCAGCAGCAGCAGCAGCAGCAGC 1327  
QY 341 ThrSerSerSerGlyThrAlaGlnArgSerLeuLeuSerProLeuHisArgSerGlySer 360  
DB 1328 ACCAGCAGCAGTGGCAGCCGCCAGAGGTCCCTTTATCTCCACTGCATCGAAGTGGAGC 1387  
QY 361 ValGluAspAsnTyrHisThrSerPheLeuPheGlyLeuSerThrSerGlnGlnHisLeu 380  
DB 1388 GTGGAGGACAATTACACACAGCTTCCTTTTGGGCTTTCACACAGCAGCAGCAGCAGC 1447  
QY 381 ThrLysSerAlaGlyLeuLysGlyTyrHisSerAspIleLeuAlaProGlnThr 400  
DB 1448 ACGAGTCTGCTGGCTGGGCTTAAAGGCTGGCAGCTCGGATATCTTGGCCCCCAGACC 1507  
QY 401 SerThrProSerLeuThrSerSerTyrPhealThrGluSerSerHisPheTyrSer 420  
DB 1508 TCTACCCCTTCCCTGACAGCAGCTGGTATTTTGGCCAGAGTCTCAGACATCTTACTCT 1567  
QY 421 AlaSerAlaIleTyrGlyGlySerAlaSerTyrSerAlaTyrSerCysSerGlnLeuPro 440  
DB 1568 GCTCAGCCATCTACGGAGGAGTGCAGTACTCTGTGCTACAGCTGACGAGCAGTCCCC 1627  
QY 441 ThrCysGlyAspGlnValTyrSerValArgArgGlnLysProSerAspArgAlaAsp 460  
DB 1628 ACTTCGAGACCAAGTCTATTCTGTGCGAGGCGGAGAGCAAGCAAGTGCAGAGCTGAC 1687  
QY 461 SerArgArgSerTyrHisGluSerProPheGluLysGlnPheLysArgArgSerCys 480  
DB 1688 TCGCGCGAGCTGGCATGAAGAGAGGCCCTTTTGAAGAGCAGTTTAAACGAGAGTGC 1747

QY 481 GlnMetGluPheGlyGluSerLleMetSerGluAsnArgSerArgGluGluLeuGlyLys 500  
Db 1748 CAATGGAAATTTGAGAGAGCATCATGTCAGAGAACAGGTCACGGGAAGAGCTGGGGAAA 1807  
QY 501 ValGlySerGlnSerSerPheSerGlySerMetGluLleIleGluValSer 517  
Db 1808 GTGGGCAGTCAGTCAGTCTTTTGGGCAGCATGGAAATCATTCAGGTCTCC 1858

RESULT 2  
AX441229 3332 bp DNA linear PAT 28-JUN-2002  
LOCUS Sequence 20 from Patent WO0226997.  
DEFINITION AX441229  
ACCESSION AX441229  
VERSION AX441229.1 GI:21665771  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
REFERENCE  
AUTHORS Luche, R.M. and Wei, B.  
TITLE Dsp-16 dual-specificity phosphatase  
JOURNAL Patent: WO 0226997-A 20 04-APR-2002;  
Ceptyr, Inc. (US)  
FEATURES  
source Location/Qualifiers  
1..3332  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Alignment Scores:  
Pred. No.: 0 Length: 3332  
Score: 517.00 Matches: 517  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-09-964-277-21 (1-517) x AX441229 (1-3332)

QY 1 MetLeuProLeuSerLeuGlnThrValPheSerLeuTyrrPheTrpValAsnTrpArg 20  
Db 842 ATGTGGCTCTCTCTCTACAGCTGTTTCTACTGTAATCTGGGTAACTGGGAAGA 901  
QY 21 AlaSerThrLeuPheThrCysLeuGlnGluMetGlnAsnGlyIleGlyTyrrVal 40  
Db 902 GCTTCAACTCTGTTCACTGCTTGCAGAGCTGATGCAGCAGATGGGATTGGTTATGTG 961  
QY 41 LeuAsnAlaSerAsnThrCysProLysProAspPheIleProGluSerHisPheLeuArg 60  
Db 962 TTAATGCCAGCAATACCTGTCCAAAGCCTGACTTTATCCCGAGTCTCATTTCTCTGCT 1021  
QY 61 ValProValAsnAspSerPheCysGluLysIleLeuProTrpLeuAspLysSerValAsp 80  
Db 1022 GTGGCTGTGAATGACAGCTTTTGAGAAATTTTGGCGTGGTTGGCAAAATCAGTAGAT 1081  
QY 81 PheLeuGluLysAlaLysAlaSerAsnGlyCysValLeuValHisCysLeuAlaGlyIle 100  
Db 1082 TTCATTGAGAAAGCAAAAGCCCTCAATGGATGTGTTCTAGTGACCTGTTTAGCTGGGATC 1141  
QY 101 SerArgSerAlaThrIleAlaIleAlaTyrrIleMetLysArgMetAspMetSerLeuAsp 120  
Db 1142 TCCCGCTCCGCACATCGCTATCGCTATCATCATGATGAGAGATGACATGCTTTAGAT 1201  
QY 121 GluAlaTyrrArgPheValLysGluLysArgProThrIleSerProAsnPheAsnPheLeu 140  
Db 1202 GAAGCTACAGATTGTGMAAGAAAAAGACCTACTATATCTCAAACTTCAATTTCTG 1261  
QY 141 GlyGlnLeuLeuAspTyrrGluLysLysIleLysAsnGlnThrGlyAlaSerGlyProLys 160  
Db 1262 GGCCAACTCTCTGGAGTATGAGAAAGAAATTAAGAACAGACTGGAGCATCAGGGCCAAAG 1321

RESULT 3

QY 161 SerLysLeuLysLeuHisLeuGluLysProAsnGluProValProAlaValSerGlu 180  
Db 1322 AGCAAACTCAAGCTGCTGCACCTGGGAAGAGCAAAATGAACCTGTCCCTGCTGTCTCAGAG 1381  
QY 181 GlyGlyGlnLysSerGluThrProLeuSerProProCysAlaAspSerAlaThrSerGlu 200  
Db 1382 GGTGGACAGAAAAGCAGAGCGCCCTCAGTCCACCCCTGTGCCGACTCTCTACCTCAGAG 1441  
QY 201 AlaAlaGlyGlnArgProValHisProAlaSerValProSerValProSerValGlnPro 220  
Db 1442 GCAGCAGGACAAAAGGCCCGGTGCATCCCGCAGCGTCCCGAGCGTCCCGAGCGCG 1501  
QY 221 SerLeuLeuGluAspSerProLeuValGlnAlaLeuSerGlyLeuHisLeuSerAlaAsp 240  
Db 1502 TCGCTGTTAGAGCAGCGCCGCTGTGTACAGGCGCTCAGTGGCTGCACCTGTCCCGCAGAG 1561  
QY 241 ArgLeuGluAspSerLeuLysLeuLysArgSerPheSerLeuAspLysSerValSer 260  
Db 1562 AGGTGGAGACAGCAATAAGCTCAAGCGTTCCTCTCTCTGATATCAAAATCAGTTTCA 1621  
QY 261 TyrSerAlaSerMetAlaAlaSerLeuHisGlyPheSerSerGluAspAlaLeuGlu 280  
Db 1622 TATTCAGCCAGCATGGCAGCATCTTACATGCTTCTCTCATCAGAGATGCTTTGNA 1681  
QY 281 TyrTyrrLysProSerThrThrLeuAspGlyThrAsnLysLeuCysGlnPheSerProVal 300  
Db 1682 TACTACAAACTTCCACTACTCTGGATGGGACCAACAGCTATGCCAGTTCTCCCTGTT 1741  
QY 301 GlnGluLeuSerGluGlnThrProGluThrSerProAspLysGluGluAlaSerIlePro 320  
Db 1742 CAGGAATATCGGAGCAGACTCCCGAAACCACTCTCTGATAAGAGAGACCCAGCATCCCC 1801  
QY 321 LysLysLeuGlnThrAlaArgProSerAspSerGlnSerLysArgLeuHisSerValArg 340  
Db 1802 AAGAGCTGCAGACCGCCAGGCCCTTCAGACACCCAGCAGCAGCGATTGCTTCGCTCAGA 1861  
QY 341 ThrSerSerSerGlyThrAlaGlnArgSerLeuLeuSerProLeuHisArgSerGlySer 360  
Db 1862 ACCAGCAGCAGTGGCCCGCCAGAGTCCCTTTTATCTCCACTGCATCGAAGTGGGAGC 1921  
QY 361 ValGluAspAsnTyrrHisThrSerPheLeuPheGlyLeuSerThrSerGlnGlnHisLeu 380  
Db 1922 GTGGAGGACAAATTCACACACAGCTTCTTTTTCGGCTTTCCACAGCAGCAGCACCTC 1981  
QY 381 ThrLysSerAlaGlyLeuGlyLeuLysGlyTyrrHisSerAspLysLeuAlaProGlnThr 400  
Db 1982 ACGAAGTCTGCTGGCCTGGGCTTAAAGGCTGGCCTCGGATATCTTGGCCCCCAGAGCC 2041  
QY 401 SerThrProSerLeuThrSerSerTrpTyrrPheAlaThrGluSerSerHisPheTyrrSer 420  
Db 2042 TCTACCCCTTCTCCCTGACAGCAGCTGGTATTTTGGCACAGAGTCTCTCACACTTCTACTCT 2101  
QY 421 AlaSerAlaIleTyrrGlySerAlaSerTyrrSerAlaTyrrSerCysSerGlnLeuPro 440  
Db 2102 GCCTCAGCCATCTACGAGGAGCAGTGCAGTACTCTGCTCTACAGTGCAGCAGCTGCC 2161  
QY 441 ThrCysGlyAspGlnValTyrrSerValArgArgArgGlnLysProSerAspArgAlaAsp 460  
Db 2162 ACTTCGGAGACCAAGTCTATTCTGTGCGCAGCGCAGAGCCAAAGTGCAGAGCTGAC 2221  
QY 461 SerArgArgSerTrpHisGluGluSerProPheGluLysGlnPheLysArgArgSerCys 480  
Db 2222 TCGCGCGGAGCTGGCATGAGAGAGCCCTTTGAAAGCAGTTTAAACGCAAGAGCTGC 2281  
QY 481 GlnMetGluPheGlyGluSerIleMetSerGluAsnArgSerArgGluLeuLeuGlyLys 500  
Db 2282 CAAATGGAATTTGGAGAGAGCATCATGTGCAGAGAACAGGTTCACGGGAAGAGCTGGGAAA 2341  
QY 501 ValGlySerGlnSerPheSerGlySerMetGluLleIleGluValSer 517  
Db 2342 GTGGGCAGTCAGTCTAGCTTTTTCGGGCAGCATGGAAATCATTTAGAGTCTCC 2392

AX260342 AX260342 1998 bp DNA linear PAT 26-OCT-2001  
LOCUS  
DEFINITION Sequence 3 from Patent WO0173059.  
ACCESSION AX260342  
VERSION AX260342.1 GI:16509305  
KEYWORDS Homo sapiens (human)  
SOURCE  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
REFERENCE  
AUTHORS Meyers,R.A.  
TITLE 38692 and 21117: dual specificity phosphatase molecules and uses  
therefor  
JOURNAL Patent: WO 0173059-A 3 04-OCT-2001;  
Millennium Pharmaceuticals, Inc. (US)  
FEATURES  
source Location/Qualifiers  
1..1998  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Alignment Scores:  
Pred. No.: 0 Length: 1998  
Score: 488.00 Matches: 488  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 94.39% Indels: 0  
DB: 6 Gaps: 0  
US-09-964-277-21 (1-517) x AX260342 (1-1998)  
QY 30 GluLeuMetGlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerAsnThrCysProLys 49  
DB 532 GAGCTGATGCACAGAAATGGGATGGTATGTTAAATGCCACATATCTGTCCAAAG 591  
QY 50 ProAspPheIleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGlu 69  
DB 592 CCGTGAATTTATCCCGAGTCTCATTTCTCGGTGCTGCTGGAATGACAGCTTTGTGAG 651  
QY 70 LysIleLeuProThrLeuAspLysSerValAspPheIleGlyLysAlaLysAlaSerAsn 89  
DB 652 AAAATTTTCCCGTGTGTGACAAATCAGTAGATTTCATTGAGAAGCAAAAGCTTCCAAAT 711  
QY 90 GlyCysValLeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAla 109  
DB 712 GGAATGTTCTAGTGCACCTGTTTACCTGGATCTCCCGCTCCGCCACCATCGCTATCGCC 771  
QY 110 TyrIleMetLysArgMetAspMetSerLeuAspGluAlaTyrArgPheValLysGluLys 129  
DB 772 TACATCATGAAGAGATGGACATGCTTTTAGATGAAGCTTACAGATTGTGAAAGAAAAA 891  
QY 130 ArgProThrIleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGluLysLys 149  
DB 832 AGACCTACTATATCTCCAAACTTCAATTTCTGGCCCACTCTCTGGACTATGAGAAGAG 891  
QY 150 IleLysAsnGlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuLeuHisLeuGlu 169  
DB 892 ATTAAGAACCACTGGAGCATCAGGGCCAAAGAGCAAACTCAAGCTGCTGCACCTGGAG 951  
QY 170 LysProAsnGluProValProAlaValSerGluGlyGlnLysSerGluThrProLeu 189  
DB 952 AGCCAAATGAACCTGTCTCTGCTCAGAGGTGGACAGAAAGCGAGAGCGCCCTC 1011  
QY 190 SerProProCysAlaAspSerAlaThrSerGluAlaAlaGlyGlnArgProValHisPro 209  
DB 1012 AGTCCACCTGTGCGGACTCTCTACCTACCTCAGAGGCAGCAGCAAAAGCGCGTGCATCCC 1071  
QY 210 AlaSerValProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuVal 229  
DB 1072 GCCAGCGTCCCGAGCTGCCAGCGTGCAGCGCTGCTGTTAGAGGACAGCCCGCTGGTA 1131

QY 230 GlnAlaLeuSerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnLysLeuLys 249  
DB 1132 CAGCGCTCAGTGGCTGCACCTGTCCGCAGACAGGCTGAAGACAGCAATAGACTCAAG 1191  
QY 250 ArgSerPheSerLeuAspIleLysSerValSerTyrSerAlaSerMetAlaAlaSerLeu 269  
DB 1192 CGTTCCTTCTCTCTGGATATCAAAATCAGTTTCATATTCAGCCAGCATGGCAGCATCCTTA 1251  
QY 270 HisGlyPheSerSerSerGluAspAlaLeuGluTyrTyrLysProSerThrThrLeuAsp 289  
DB 1252 CATGCTTCTCTCATCAGAAGATGCTTTGGAATACATAAAACCTTCCACTACTCTGGAT 1311  
QY 290 GlyThrAsnLysLeuCysGlnPheSerProValGlnGlnLeuSerGluGlnThrProGlu 309  
DB 1312 GGGACCAACAGCTATGCCAGTTCTCCCTGTTCCAGAACTATCGGAGCAGACTCCCGAA 1371  
QY 310 ThrSerProAspLysGluGluAlaSerIleProLysLysLeuGlnThrAlaArgProSer 329  
DB 1372 ACCAGTCTCTGATAAGAGGAGGAGCCAGCATCCCAAGAAGCTGCAGACCCGCGAGCCTCA 1431  
QY 330 AspSerGlnSerLysArgLeuHisSerValArgThrSerSerGlyThrAlaGlnArg 349  
DB 1432 GACAGCCAGAGCAAGCGATTGCAATTCGCTCAGAACCCAGCAGAGTGGACCCGCGAGAG 1491  
QY 350 SerLeuLeuSerProLeuHisArgSerGlySerValGluAspAsnTyrHisThrSerPhe 369  
DB 1492 TCCCTTTTATCTCCACTGCATCGAAGTGGAGCGTGGAGGACAATACACACAGCTTC 1551  
QY 370 LeuPheGlyLeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLys 389  
DB 1552 CTTTTCGGCCTTTCCACAGCCAGCAGCACCTCAGAAAGTCTGCTGGCTGGGCCCTTAAG 1611  
QY 390 GlyThrHisSerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTrp 409  
DB 1612 GGTGGCACTCGGATATCTTGGCCCCCAGACCTCTACCCCTTCCCTGACCCAGCAGCTGG 1671  
QY 410 TyrPheAlaThrGluSerSerHisPheTyrSerAlaSerAlaIleTyrGlyGlySerAla 429  
DB 1672 TATTTTGCACAGAGTCTCTCACACTTCTACTCTGCTCAGCCATCTACCGAGCAGTGCC 1731  
QY 430 SerTyrSerAlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerVal 449  
DB 1732 AGTTACTCTGCTACAGTCAGCCAGCTGCCCACTTGGGAGACCAAGTCTATTCTCTGG 1791  
QY 450 ArgArgGlnLysProSerAspArgAlaAspSerArgSerTrpHisGluGluSer 469  
DB 1792 CGCAGCGCGCAAGAGCCAAAGTGACAGAGCTGACTCGCGCGAGCTGGCATGAAGAGAG 1851  
QY 470 PropheGluLysGlnPheLysArgSerCysGlnMetGluPheGlyGluSerIleMet 489  
DB 1852 CCCTTTGAAAAGCAGTTTAAACGCAAGCTGCCAAATGGAATTTGGAGAGCATCATG 1911  
QY 490 SerGluAsnArgSerArgGluGluLeuGlyLysValGlySerGlnSerSerPheSerGly 509  
DB 1912 TCAGAGAACAGTCACTCGGGAAGAGCTGGGGAAGAGTGGCAGTCAGTCTAGCTTTCCGGC 1971  
QY 510 SerMetGluIleIleGluValSer 517  
DB 1972 AGCATGGAATCATTTAGGTCTCC 1995  
RESULT 4  
AX713989  
LOCUS AX713989 2102 bp DNA linear PAT 15-APR-2003  
DEFINITION Sequence 673 from Patent EP1293569.  
ACCESSION AX713989  
VERSION AX713989.1 GI:29888917  
KEYWORDS Homo sapiens (human)  
SOURCE  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
REFERENCE  
AUTHORS Isogai,T., Sugiyama,T., Otsuki,T., Wakamatsu,A., Sato,H., Ishii,S.,



Yamamoto, J.I., Isono, Y., Hio, Y., Otsuka, K., Nagai, K., Irie, R.,  
Tamechika, I., Seki, N., Yoshikawa, T., Otsuka, M., Nagahari, K. and  
Masuho, Y.

TITLE  
JOURNAL  
Full-length cDNAs  
Patent: EP 1293569-A 673 19-MAR-2003;  
Helix Research Institute (JP) ; Research Association for  
Biotechnology (JP)

FEATURES  
source  
Location/Qualifiers  
1..2102  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

## ORIGIN

## Alignment Scores:

Pred. No.: 0 Length: 2102  
Score: 488.00 Matches: 488  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 94.39% Indels: 0  
DB: 6 Gaps: 0

US-09-964-277-21 (1-517) x AX713989 (1-2102)

QY 30 GluLeuMetGlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerAsnThrCysProLys 49  
Db 587 GAGCTGATGCAGCAGATGGGATTGGTTATGTTAAATGCCAATACTGTGCCAAG 646  
QY 50 ProAspPheLeuProGluSerHisPheLeuArgValProValAsnAspSerPheCysGlu 69  
Db 647 CTTGACTTTATCCCCGAGTCTCAATTCCTGCTGGCTGCTGGAATGACAGCTTTTGAG 706  
QY 70 LysIleLeuProTrpLeuAspLysSerValAspPheIleGluLysAlaLysAlaSerAsn 89  
Db 707 AAAATTTGGCGTGTGGACAAATCACTAGATTTCATTGAGAAAGCAAAAGCTCCAT 766  
QY 90 GlyCysValLeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAla 109  
Db 767 GGATGTGTTCTAGTGCATCTGTTAGCTGGGACTCTCCGCTCCGACCATCGTATGCC 826  
QY 110 TyrIleMetLysArgMetAspMetSerLeuAspGluAlaTyrArgPheValLysGluLys 129  
Db 827 TACATCATGAGAGAGATGGACATGTCCTTTAGATGAGCTTACAGATTGTGGAAGAAAA 886  
QY 130 ArgProThrIleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGluLysLys 149  
Db 887 AGACCTACTATATCTCCAAACTTCAATTTCTGGCCAACTCTCTGGCACTATGAGAAGAG 946  
QY 150 IleLysAsnGlnThrGlyAlaSerClyProLysSerLysLysLeuLysLeuHisLeuGlu 159  
Db 947 ATTAAAGAACAGACTGGAGCATCAGGGCCAAAGACCAAACTCAAGCTGCTGCACCTGGAG 1006  
QY 170 LysProAsnGluProValProAlaValSerGluGlyGlyGlnLysSerGluThrProLeu 189  
Db 1007 AAGCCAAATGAACCTGTCTCTGCTGTCTCAGAGGTGGACAGAAACGAGAGCGCCCTC 1066  
QY 190 SerProProCysAlaAspSerAlaThrSerGluAlaGlyGlnArgProValHisPro 209  
Db 1067 AGTCCACCTGTGTCCGACTCTGTACTCTCAGAGCAGCAGCAAAAGCCCGTGCATCCC 1126  
QY 210 AlaSerValProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuVal 229  
Db 1127 GCCACGTCGCCAGCGTCCAGCGTGCAGCGCTGCTGCTGTGTAGAGGACGCCCGCTGGA 1186  
QY 230 GlnAlaLeuSerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnLysLeuLys 249  
Db 1187 CAGCGGCTCAGTGGCTGCCTGTCCGACAGAGGTGGAGACCAATAAGCTCAAG 1246  
QY 250 ArgSerPheSerLeuAspIleLysSerValSerTyrSerAlaSerMetAlaAlaSerLeu 269  
Db 1247 CTTCTCTCTCTCTGGATATCAATCATCTGTTTCAATTCAGCCAGCATGGCAGCATCTTA 1306  
QY 270 HisGlyPheSerSerSerGluAspAlaLeuGluTyrTyrLysProSerThrThrLeuAsp 289

Db 1307 CATGCTTCTCTCTCATCAGAAGATGCTTTGGAATPACTACAAACCTTCCACTACTCTGGAT 1366  
QY 230 GlyThrAsnLysLeuCysGlnPheSerProValGlnGluLeuSerGluGlnThrProGlu 309  
Db 1367 GGGACCAACAGCTATGCCAGTTCTCCCTGTTCAGGAACCTATCGAGCAGACTCCCGAA 1426  
QY 310 ThrSerProAspLysGluGluAlaSerIleProLysLysLeuGlnThrAlaArgProSer 329  
Db 1427 ACCAGTCTGTATAGGAGGAAGCCAGCATCCCCAAGAAGCTGCAGACTGCCAGGCTTCA 1486  
QY 330 AspSerGlnSerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArg 349  
Db 1487 GACAGCCAGACAGCGATTGCTTCGTCAGAACCCAGCAGCAGTGGCACCAGCCAGAGG 1546  
QY 350 SerLeuLeuSerProLeuHisArgSerGlySerValGluAspAsnTyrHisThrSerPhe 369  
Db 1547 TCCCTTTTATCTCCACTGCATCGAAGTGGAGAGGTGGAGACAATTACCACACAGCTTC 1606  
QY 370 LeuPheGlyLeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLeuLys 389  
Db 1607 CTTTTGGCCCTTCCACGCCAGCAGCAGCAGCTCAGGAAGTCTGCTGGCTGGCCCTTAAG 1666  
QY 390 GlyTrpHisSerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTrp 409  
Db 1667 GGTGGCACTCGGATATCTTTGGCCCCCAGACCTCTACCCCTTCCCTGACCCAGCTGG 1726  
QY 410 TyrPheAlaThrGluSerSerHisPheTyrSerAlaSerAlaIleTyrGlyGlySerAla 429  
Db 1727 TATTTGGCCAGAGTCTCTCACACTTCTCTGCTCAGCCATCTACGGAGCAGTGCC 1786  
QY 430 SerTyrSerAlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerVal 449  
Db 1787 AGTTACTCTGCTACAGCTGCAGCCAGCTCCCACTTGGGAGACCAAGTCTATTCTGTG 1846  
QY 450 ArgArgArgGlnLysProSerAspArgAlaAspSerArgArgSerTrpHisGluGluSer 469  
Db 1847 CGCAGGGCGCAGAGCAAGTGCAGAGCTGACTCGCGCGAGCTGGCATGAGAGAGC 1906  
QY 470 ProPheGluLysGlnPheLysArgArgSerCysGlnMetGluPheGlyGluSerIleMet 489  
Db 1907 CCCTTTGAAGCAGTTTAAACGCAAGAGCTGCCAATTGGAAATTTGGAGAGACATCATG 1966  
QY 490 SerGluAsnArgSerArgGluGluLeuGlyLysValGlySerGlnSerSerPheSerGly 509  
Db 1967 TCAGAGAACAGTCAACGGGAAGAGCTGGGAAAAGTGGCAGTCACTAGCTTTTCGGGC 2026  
QY 510 SerMetGluIleIleGluValSer 517  
Db 2027 AGCATGGAATCATTGAGGCTCC 2050  
RESULT 5  
AK055973 2102 bp mRNA linear PRI 01-AUG-2002  
LOCUS  
DEFINITION  
Homo sapiens cDNA FLJ31411, clone NT2NE2000214, moderately  
similar to DUAL SPECIFICITY PROTEIN PHOSPHATASE 8 (EC 3.1.3.48).  
ACCESSION  
AK055973.1 GI:16550835  
VERSION  
oligo capping; fis (full insert sequence).  
KEYWORDS  
Homo sapiens (human)  
SOURCE  
Homo sapiens  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1  
AUTHORS  
Tashiro, H., Yamazaki, M., Watanabe, K., Kumagai, A., Itakura, S.,  
Fukuzumi, Y., Fujimori, Y., Komiyama, M., Sugiyama, T., Irie, R.,  
Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y.,  
Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H.,  
Masuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K.,  
Wagatsuma, M., Murakawa, K., Kaneshori, K., Takahashi-Fujii, A.,  
Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S.,  
Nagahari, K., Masuho, Y., Nagai, K. and Isogai, T.  
NEDO human cDNA sequencing project

TITLE

Unpublished  
 REFERENCE 2 (bases 1 to 2102)  
 AUTHORS Isogai,T., Otsuki,T. and Sugiyama,T.  
 TITLE Direct Submission  
 JOURNAL Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute,  
 Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
 (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)  
 NEDO human cDNA sequencing project supported by Ministry of  
 Economy, Trade and Industry of Japan; cDNA full insert sequencing:  
 Research Association for Biotechnology (RAB); cDNA library  
 construction: Helix Research Institute (HRI) (supported by Japan  
 Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,  
 HRI, and Biotechnology Center, National Institute of Technology and  
 Evaluation; clone selection for full insert sequencing: RAB and  
 HRI.

FEATURES  
 source Location/Qualifiers  
 1..2102  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="NT2NE200214"  
 /cell\_line="NT2"  
 /cell\_type="teratocarcinoma"  
 /clone\_lib="NT2NE2"  
 /notes="cloning vector: pME18SFL3-mRNA from NT2 neuron  
 after the differentiation of NT2 neuronal precursor  
 cells."  
 56..2053  
 /notes="unamed protein product"  
 /codon\_start=1  
 /protein\_id="BAB71060.1"  
 /db\_xref="GI:16550836"  
 /translation="MAHEMIGQIVTERLVALLESGETKVLILIDSRPFVYNTSHILE  
 AINNGSLKMRLOQKVLIELIQAQAKHVDICQKVVYDQSQDVSLSDDC  
 FLVLGLKLEKSFNSVHLLAGFAFSCFGLCEGKSLVPTCTISQPCLPVANGPT  
 LILPNLYLQCDRLVNLKMQNGIGYVLNASNTCPKDFIPESHFRFVPWDSFCBK  
 LIPWLDKVFLEKAKSNGCIVLHLAGISRSATIAIATMKRMDMSLDEAYRFVK  
 KPTISPNFNLQGLLDYEKKIKNOTGASGPKSLKLLHLEKPNPEVPVAVSEGGQKSE  
 TPLSPCADSATSEAGQRPVSPVSPSLLEDPLVQALSGHLHLSADRLBD  
 SNLKRKSFSLDKSVYSASMAASHLFFSSSDALEYKPKSTLDGKLCQFSPVQE  
 LSPQTPETSPDKEEASIPKLOTARPSQSKRLHVSRTSSGTAQRLSLSPILHRSQS  
 VEDNYHTSLFLGSLTQSHLTKSAGLKGKHSIDILAPQTSPLTSSWYFATSSHF  
 YNSAIIYGSASYSAYSCQLPTGDDQYVYVRRRKPSDRADRSRHSRHSRPFKEQKF  
 RRSQCFEGESIMSENRSREELKVGQSQSFSGSMELIEVS"

CDS  
 56..2053  
 /notes="unamed protein product"  
 /codon\_start=1  
 /protein\_id="BAB71060.1"  
 /db\_xref="GI:16550836"  
 /translation="MAHEMIGQIVTERLVALLESGETKVLILIDSRPFVYNTSHILE  
 AINNGSLKMRLOQKVLIELIQAQAKHVDICQKVVYDQSQDVSLSDDC  
 FLVLGLKLEKSFNSVHLLAGFAFSCFGLCEGKSLVPTCTISQPCLPVANGPT  
 LILPNLYLQCDRLVNLKMQNGIGYVLNASNTCPKDFIPESHFRFVPWDSFCBK  
 LIPWLDKVFLEKAKSNGCIVLHLAGISRSATIAIATMKRMDMSLDEAYRFVK  
 KPTISPNFNLQGLLDYEKKIKNOTGASGPKSLKLLHLEKPNPEVPVAVSEGGQKSE  
 TPLSPCADSATSEAGQRPVSPVSPSLLEDPLVQALSGHLHLSADRLBD  
 SNLKRKSFSLDKSVYSASMAASHLFFSSSDALEYKPKSTLDGKLCQFSPVQE  
 LSPQTPETSPDKEEASIPKLOTARPSQSKRLHVSRTSSGTAQRLSLSPILHRSQS  
 VEDNYHTSLFLGSLTQSHLTKSAGLKGKHSIDILAPQTSPLTSSWYFATSSHF  
 YNSAIIYGSASYSAYSCQLPTGDDQYVYVRRRKPSDRADRSRHSRHSRPFKEQKF  
 RRSQCFEGESIMSENRSREELKVGQSQSFSGSMELIEVS"

ORIGIN  
 Alignment Scores:  
 Pred. No.: 0 Length: 2102  
 Score: 488.00 Matches: 488  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 94.39% Indels: 0  
 DB: 9 Gaps: 0

US-09-964-277-21 (1-517) x AK055973 (1-2102)

Qy 30 GluLeuMetGlnGlnAnGlyIleGlyTyrValLeuAsnAlaSerAsnThrCysProlys 49  
 Db 587 GAGCTGATGACAGAGATGGATGGTATGTTGTTAAATGCCAGCAATACCTGTCCAAAG 646  
 Qy 50 ProAspPheIleProGluSerHisPheLeuAqValProValAsnAspSerPheCysGlu 69  
 Db 647 CCTGACCTTATCCCGAGTCTGATTTCTCGCGTGGCTGTGAATGACAGCTTTGTGAG 706  
 Qy 70 LysIleLeuProTrpLeuAspLysSerValAppPheIleGlyAlaLysAlaSerAsn 89  
 Db 707 AAAATTTTGGCGTGGTGGCAAAATCAGTAGATTTTCAATTCAGAAAGCAAAAGCCCTCCCAAT 766  
 Qy 90 GlyCysValLeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAla 109  
 Db 767 GGATGTGTTCTAGTGACCTGTTTAGCTGGGATCTCCCGCTCCGCCACCATGCTATCGCC 826  
 Qy 110 TyrIleMetLysArgMetAspMetSerLeuAspGluAlaTyrArgPheValLysGluLys 129

Db 827 TACATCATGAGAGAGATGGACATGCTTTTAGATGAGCTTACAGATTTGTGAAAGAAAA 886  
 Qy 130 ArgProThrIleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGlyLys 149  
 Db 887 AGACCTACTATATCTCCAACTTCAATTTTCTGGGCAACTCTCTGGACTATGAGAAG 946  
 Qy 150 IleLysAsnGlnThrGlyAlaSerGlyProLysSerLysLeuLeuHisLeuGlu 169  
 Db 947 ATTAAGAACCCAGACTGGAGCATCAGGGCCAAAGAGCAAACTCAAGCTGCTGCACCTGGAG 1006  
 Qy 170 LysProAsnGluProValProAlaValSerGluGlyGlnLysSerGluThrProLeu 189  
 Db 1007 AGCCCAATGAACCTGCTCCTGCTCAGAGGGTGGACAGAAAGCGAGAGCCCTC 1066  
 Qy 190 SerProCysAlaAspSerAlaThrSerGluAlaAlaGlyGlnArgProValHisPro 209  
 Db 1067 AGTCCACCCTGTGGGACTCTGCTACCTCAGAGCAGCAGGACCAAGCCCGTGCATCCC 1126  
 Qy 210 AlaSerValProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuVal 229  
 Db 1127 GCGAGGTGCCAGGCTGCCAGCGTGCAGCGCTGCTGTAGAGGACAGCCCGCTGGTA 1186  
 Qy 230 GlnAlaLeuSerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnLysLeuLys 249  
 Db 1187 CAGGCGCTCAGTGGGCTGCACCTGTCGACAGCAGGCTGGAAAGACAGCAATAAGCTCAAG 1246  
 Qy 250 ArgSerPheSerLeuAspIleLysSerValSerTyrSerAlaSerMetAlaAlaSerLeu 269  
 Db 1247 CGTTCCTTCTCTCGATATCAATCAGTTTCATATCAGCCAGCATGGCAGCATCTT 1306  
 Qy 270 HisGlyPheSerSerSerGluAspAlaLeuGluTyrTyrLysProSerThrThrLeuAsp 289  
 Db 1307 CATGGCTCTCTCATAGAAGATGCTTTGGAATACTACAAACCTTCACCTACTCTCGAT 1366  
 Qy 290 GlyThrAsnLysLeuCysGlnPheSerProValGlnGluLeuSerGluGlnThrProGlu 309  
 Db 1367 GGGACCAACAGCTATGCGAGTTCCTCCCTGTTCCAGAACTATCGGAGCAGACTCCCGAA 1426  
 Qy 310 ThrSerProAspLysGluGluAlaSerIleProLysLysLeuGlnThrAlaArgProSer 329  
 Db 1427 ACCAGTCTCTGATAAGAGAGAGCCAGCATCCCCAAGAGCTGCAGACTGCCAGGCCCTTCA 1486  
 Qy 330 AspSerGlnSerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArg 349  
 Db 1487 GACGCCAGCAGCAGCATTCGATTCGGTCCAGACCAGCAGCAGTGGCAGCCGCCAGG 1546  
 Qy 350 SerLeuLeuSerProLeuHisArgSerGlySerValGluAspAsnTyrHisThrSerPhe 369  
 Db 1547 TCCTTTTATCTCCACTGCATCGAAGTGGAGCGTGGAGGACATTTACCACACAGCTTC 1606  
 Qy 370 LeuPheGlyLeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLys 389  
 Db 1607 CTTTTCGGCTTTTCCACCAGCAGCAGCAGCATCCGAAAGTCTGCTGGCCTTGAAG 1666  
 Qy 390 GlyTrpHisSerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTip 409  
 Db 1667 GCGTGCATCTCGATATCTTGGCCCCCAGACCTCTACCCCTTCCTGACCCAGCAGCTGG 1726  
 Qy 410 TyrPheAlaThrGluSerSerHisPheThrSerAlaSerAlaIleTyrGlyGlySerAla 429  
 Db 1727 TATTTTGGCAGAGTCTCCTACCTTCTCTGCTCAGCCATCTACGAGGAGCAGTGCC 1786  
 Qy 430 SerTyrSerAlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerVal 449  
 Db 1787 AGTACTCTGCTACAGCTGCAGCCAGCTGCCCTTGGCGAGACCAAGTCTATTCTGTG 1846  
 Qy 450 ArgArgGlnLysProSerAspArgAlaAspSerArgArgSerTrpHisGluGluSer 469  
 Db 1847 CGCAGGCGCAGAACCAAGTGACAGACTGACTCGCGCGAGCTGCGCATGAAGAGAGC 1906  
 Qy 470 ProPheGluLysGlnPheLysArgArgSerCysGlnMetGluPheGlyGluSerIleMet 489

```
Db 1907 CCCTTTGAAAAGCAGTTTAAACGACGAGCTGCCAAATGGAATTTGGAGAGACATCATG 1966
Qy 490 SerGluAsnArgSerArgGluGluLeuGlyValGlySerGlnSerPheSerGly 509
Db 1967 TCAGAGAACAGCTCACGGGAAGAGCTGGGAAAGTGGCCAGTCAGTCTAGCTTTTCGGGC 2026
Qy 510 SerMetGluLeuLeuGluValSer 517
Db 2027 AGCATGGAATCATTGAGGTCTCC 2050

RESULT 6
AXI80875 2732 bp DNA linear PAT 06-AUG-2001
LOCUS AXI80875 Sequence 2 from Patent WO0146394.
DEFINITION AXI80875
ACCESSION AXI80875
VERSION AXI80875.1 GI:15132703
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Plowman,G.D., Martinez,R., Whyte,D., Manning,G., Sudarsanam,S.,
Hill,R.J. and Planagan,P.
TITLE Mammalian protein phosphatases
JOURNAL Patent: WO 0146394-A 2 28-JUN-2001;
Sugen, Inc. (US)
FEATURES
source
1..2732
Location/Qualifiers
1..2732
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 2732
Score: 488.00 Matches: 488
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 94.39% Indels: 0
DBs: 6 Gaps: 0

US-09-964-277-21 (1-517) x AXI80875 (1-2732)
Qy 30 GluLeuMetGlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerAsnThrCysProlys 49
Db 1069 GAGCTGATGACGAGATGGGATTGGTTATGTGTTAAATGCCAGCAATACCTGTCCAAAG 1128
Qy 50 ProAspPheLeuProGluSerHisPheLeuArgValProValAsnAspSerPheCysGlu 69
Db 1129 CCTGACTTTATCCCGAGTCTCATTTCCCTCGGTGCTGTAATGACAGCTTTGTGAG 1188
Qy 70 LysIleLeuProTrpLeuAspLysSerValAspPheIleGluLysAlaLysAlaSerAsn 89
Db 1189 AAATTTTCGGTGGTGGCAATCAGTAGATTTCATTGAGAAACAAAGCCCTCCAT 1248
Qy 90 GlyCysValLeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAla 109
Db 1249 GGATGTGTTCTAGTGCACTGTTTAGCTGGGATCTCCCGCTCCGCCACCATCGCTATCGCC 1308
Qy 110 TyrIleMetLysArgMetAspMetSerLeuAspGluAlaTyrArgPheValLysGluLys 129
Db 1309 TACATCATGAGAGATGGACATGCTTTAGATGAGCTTACAGATTGTGAAAGAAAAA 1368
Qy 130 ArgProThrIleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGlyLysLys 149
Db 1369 AGACCTACTATATCTCCAACTTCATTTTCCTGGCCAACTCCCTGGACTATGAGAAGAG 1428
Qy 150 IleLysAsnGlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuHisLeuGlu 169
Db 1429 ATTAAGAACACAGCTGAGCATCAGGCCCAAGAGCAAACTCAAGCTGTGCACCTGGAG 1488
Qy 170 LysProAsnGluProValProAlaValSerGluGlyGlyGlnLysSerGluThrProLeu 189
```

RESULT 7  
AX278461

```
Db 1489 AAGCAAATGAACCTGCTCCCTGCTGCTCAGAGGGTGGACAGAAAAAGCAGAGCCCTTC 1548
Qy 190 SerProCysAlaAspSerAlaThrSerGluAlaAlaGlyGlnArgProValHisPro 209
Db 1549 AGTCCACCTGTGCCGACTCTGCTACTCTCAGAGGACAGGACAAAGGCCCTGTCATCCC 1608
Qy 210 AlaSerValProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuVal 229
Db 1609 GCCAGCTGCCAGCGTCCGACGCTGCAGCCCTCCTGTTAGAGACAGCCCGCTGGTA 1668
Qy 230 GlnAlaLeuSerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnLysLys 249
Db 1669 CAGCGCTCAGTGGGTGACCTGTCCGACAGACAGCTGGAGACAGCAATTAAGCTCAAG 1728
Qy 250 ArgSerPheSerLeuAspLysSerValSerTyrSerAlaSerMetAlaLysLeu 269
Db 1729 CGTTCCTTCCTCTGGATATCAAAATCAGTTTCATATTCAGCCAGCATGCGAGCATCTTA 1788
Qy 270 HisGlyPheSerSerGluAspAlaLeuGluTyrTyrLysProSerThrThrLeuAsp 289
Db 1789 CATGGCTTCCTCATCAGAAAGATGCTTTGGAAATCTACAAACCTTCCACTACTCTGGAT 1849
Qy 290 GlyThrAsnLysLeuCysGlnPheSerProValGlnGluLeuSerGluGlnThrProGlu 309
Db 1849 GGGACCAACAAGCTATGCCAGTTCTCCCTGTTTCAGAAACTATCGGAGCAGACTCCCGAA 1908
Qy 310 ThrSerProAspLysGluGluAlaSerIleProLysLysLeuGlnThrAlaArgProSer 329
Db 1909 ACCAGTCTCTAAGAGGAGGAGCAGCATCCCAAGAAAGCTCAGACTGCCAGGCTTCA 1968
Qy 330 AspSerGlnSerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArg 349
Db 1969 GACAGCCAGACGAGCGATTGCAATTCGGTCAAGACCAGCAGCTGGCACCCGCCAGG 2028
Qy 350 SerLeuLeuSerProLeuHisArgSerGlySerValGluAspAsnTyrHisThrSerPhe 369
Db 2029 TCCCTTTTATCTCCATCTGCATCGAGTGGGAGCGTGGAGGACAATTAACACACAGCTTC 2088
Qy 370 LeuPheGlyLeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLys 389
Db 2089 CTTTTCGGCCTTTCCACCGCCAGCAGCACCTCAAGAGTCTGTGGCCTGGGCTTTAAG 2148
Qy 390 GlyTrpHisSerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTrp 409
Db 2149 GGCTGGCACTCGGATATCTGGCCCCCAGACCTCTACCCCTTCCCTGACACGAGCTGG 2208
Qy 410 TyrPheAlaThrGluSerSerHisPheTyrSerAlaSerAlaIleTyrGlySerAla 429
Db 2209 TATTTGGCCACAGAGTCCCTCACACTTCTACTCTGGCTCAGCCATCTACGGAGGCGAGTGC 2268
Qy 430 SerTyrSerAlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerVal 449
Db 2269 AGTTACTCTGCTTACAGCTGCAGCAGCTGCCACTTGGCGAGACCAAGTCTATTCTGTG 2328
Qy 450 ArgArgArgGlnLysProSerAspArgAlaAspSerArgArgSerTrpHisGluGluSer 469
Db 2329 CGCAGGCGGAGAGGCAAGTGCAGAGCTGACTCGCGCGAGCTGGCATGAAGAGAGC 2388
Qy 470 ProPheGluLysGlnPheLysArgArgSerCysGlnMetGluPheGlyGluSerIleMet 489
Db 2389 CCCTTTGAAAGACGTTTAAACGACAGAGCTGCCAAATGGAAATTTGGAGAGAGCATCATG 2448
Qy 490 SerGluAsnArgSerArgGluGluLeuGlyLysValGlySerGlnSerSerPheSerGly 509
Db 2449 TCAGAGAACAGGTACGCGGAAGAGCTGGGGAAGTGGGAGTGGGAGTCTAGCTTTTCGGGC 2508
Qy 510 SerMetGluLeuLeuGluValSer 517
Db 2509 AGCATGGAATCATTGAGGTCTCC 2532
```

LOCUS AX278461 3059 bp DNA linear PAT 02-NOV-2000  
 DEFINITION Sequence 1 from Patent WO0177340.  
 ACCESSION AX278461  
 VERSION AX278461.1 GI:16605915  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1  
 REFERENCE Duecker, K.  
 AUTHORS Identification of a dual specificity phosphatase: dusp-10  
 TITLE Patent: WO 0177340-A 1 18-OCT-2001;  
 JOURNAL MERCK PATENT GmbH (DE)  
 FEATURES  
 source Location/Qualifiers  
 1..3059  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"  
 127..2124  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="CAD10545.1"  
 /db\_xref="GI:16605916"  
 /db\_xref="REMBL:CAD10545"  
 /translation="MAHEMIGTQIVTERLVALLEGTEKVLIDSRFPVEYNTSHILE  
 AINCSKLMKRRLQDDKVLITELIOHSAKHKVDICQKVVYDOSQDVASLSDC  
 PLVLVGLKLRKSNVHLLAGAFESRCPGLCEGKSTLYPTCTISQPCLPVANIGPT  
 RILPNLYGQRDVLNKLMOQNGIYVNLASNTCPKPDFIPESHFLRVPNVNSPEK  
 LKPMDSVDFIIRAKASNGCVLHCLAGISRSATIAIYIMKRMDSLDEAIRFYKE  
 ILPTISFNFLQDLIDYEKKIKNOTGASGPKSLKLHLKPNPEVPVAVSEGGQKE  
 TPLSPACADATASQAPRHPASVPSVPSLLEDSPLVQALSGLSADRLIED  
 SNKLKRFSDIDKISVYSASMAASLHGFSSDEALEYVKPTTLDGINKLCPSPVQ  
 LSEDTPTSPDKKEASTPKKLOTARPDSQSKRLHSVPTSSGTAQRLSLPLHRSG  
 VEDNHYTHFLFGTISQOHLTKESAGLGLKWHSDILAPOTSPSLTSWVFATESHP  
 YSAIAIYGGASASVATSCQLPTCGQVYSVRRQKPSDRADRSRSHRESPEKQKQK  
 RRSQMEFGESIMSENRSREELKVGSGSPSGSMELIEVS"

ORIGIN

Alignment Scores:	0	Length:	3059
Pred. No.:	488.00	Matches:	488
Score:	100.00%	Conservative:	0
Percent Similarity:	100.00%	Mismatches:	0
Best Local Similarity:	100.00%	Indels:	0
Query Match:	94.39%	Gaps:	0
DB:	6		

US-09-964-277-21 (1-517) x AX278461 (1-3059)

QY	30	GLuLeuMetGlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerAsnThrCysProlys	49
Db	658	GAGCTGATGCAGCAGAAATGGGATTGGTTATGTTAAATGCCAGCAATACCTGTGCCAAG	717
QY	50	ProAspPheIleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGlu	69
Db	718	CTTGACTTTATCCCGGAGTCTCATTTCTCGGTGTCCTGTGAATGACAGCTTTTGTGAG	777
QY	70	LysIleLeuProTrpLeuAspLysSerValAspPheIleGluLysAlaLysAlaSerAsn	89
Db	778	AAATATTTTGGCGTGTTGGACAAATCAGTAGATTTCATTGAGAAAGCAAAAGCCTCCAAT	837
QY	90	GlyCysValLeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAla	109
Db	838	GGATGTGTTCTAGTGCATGTTTACGTGGGATCTCCGCTCCGCCACCATCGCATCGCC	897
QY	110	TyrIleMetLysArgMetAspMetSerLeuAspGluAlaTyrArgPheValLysGluLys	129
Db	898	TACATCATGAAGAGATGGACATGCTCTTAGATGAAGCTTACAGATTTGTGAAAGAAAA	957
QY	130	ArgProThrIleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGluLysIys	149
Db	958	AGACCTACTACTATCTCAAATCTCAATTTTCTGGGCCAACTCTCTGGATCATGAGAAG	1017

QY	150	IleLysAsnGlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuLeuHisLeuGlu	169
DB	1018	ATTAAAGAACAGACTGGAGCATCAGGGCCAAAGACAAACTCAAGCTCTGCACCTGGAG	1077
QY	170	LysProAsnGluProValProAlaValSerGluGlyGlyGlnLysSerGluThrProLeu	189
DB	1078	AAGCAAATGAACCTGTCCCTGTGCTCAGAGGGTGGACAGAAAAGCAGACGCCCTC	1137
QY	190	SerProProCysAlaAspSerAlaThrSerGluAlaAlaGlyGlnArgProValHisPro	209
DB	1138	AGTCACCCCTGTGCCGACTCTGCTACCTCAGAGGAGCAGGACAAAGCCCGTGATCCC	1197
QY	210	AlaSerValProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuVal	229
DB	1198	GCCAGCTGGCCAGCTGCCCCAGCTGCAGCCGTGCTGTATTAGGACAGCCCGCTGGTA	1257
QY	230	GlnAlaLeuSerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnLysLeuLys	249
DB	1258	CAGCGGTCTAGTGGGTGCACCTGTCCGCAGACAGCGCTGGAGACAGCAATAAAGCTCAAG	1317
QY	250	ArgSerPheSerLeuAspIleLysSerValSerTyrsSerAlaSerMetAlaAlaSerLeu	269
DB	1318	CGTTCCCTTCTCTGGATATCAATCAGTTTCAATTTAGCCAGCATGGCAGCATCTCTTA	1377
QY	270	HisGlyPheSerSerGluAspAlaLeuGluTyTyTyLysProSerThrThrLeuAsp	289
DB	1378	CATGGCTTCTCCTCATCAGAAAGATGTTTGGAAATACTACAACCTTCCACTACTCGAT	1437
QY	290	GlyThrAsnLysLeuCysGlnPheSerProValGlnGluLeuSerGlnGlnThrProGlu	309
DB	1438	GGGACCAACAAAGCTATGCCAGTTCTCCCTGTTCAGGAACATATCGGACAGACTCCCGAA	1497
QY	310	ThrSerProAspLysGluGluAlaSerIleProLysLysLeuGlnThrAlaArgProSer	329
DB	1498	ACCACTCCTGATAAGGGAGAGCCAGCATCCCCAAGAAGCTGCAGACCCGAGCGCTTCA	1557
QY	330	AspSerGlnSerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArg	349
DB	1558	GACAGCCAGAGCAAGCGATTGCAATTCGGTCAAGAACCAGCAGCTGGCCGCCCCAGAGG	1617
QY	350	SerLeuLeuSerProLeuHisArgSerGlySerValGluAspAsnTyHisThrSerPhe	369
DB	1618	TCCCTTTTATCTCCACTGCATCGAAGTGGAGCGTGGAGGACAATTACCACACCGACTTC	1677
QY	370	LeuPheGlyLeuSerThrSerGlnHisLeuThrLysSerAlaGlyLeuGlyLeuLys	389
DB	1678	CTTTTCGGCTTTTCCACAGCCAGCAGCACCTCAGAAAGTCTGTCTGGCTCGGGCTTAAG	1737
QY	390	GlyTrpHisSerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTrp	409
DB	1738	GGCTGGCACTCGGANATCTTGGCCCCCCAGACCTCTACCCCTTCCCTGACCCAGCAGCTGG	1797
QY	410	TyrPheAlaThrGluSerSerHisPheTyrsSerAlaSerAlaIleTyrglyGlySerAla	429
DB	1798	TATTTTGCCACAGAGTCTCACACTTCTACTCTGCTCAGCCATCTACGGAGGAGTGC	1857
QY	430	SerTyrsSerAlaTyrsSerCysSerGlnLeuProThrCysGlyAspGlnValTyrsSerVal	449
DB	1858	AGTTTACTCTGCCCTACAGCTGCACCGAGCTGCCACTTCCGAGACCAAGTCTATTCTGTG	1917
QY	450	ArgArgArgGlnLysProSerAspArgAlaAspSerArgArgSerTrpHisGluGluSer	469
DB	1918	GCGAGCGCGCAGAAGCCAGTGACAGAGCTGACTCGCGCGGAGCTGCATGAAGAGAGC	1977
QY	470	ProPheGluLysGlnPheLysArgArgSerCysGlnMetGluPheGlyGluSerIleMet	489
DB	1978	CCCTTTGAAAAGCAGTTTAAACCCAGAAGCTCCCAAAATGGAAATTTGGAGAGAGCATCATG	2037
QY	490	SerGluAsnArgSerArgGluGluLeuGlyLysValGlySerGlnSerSerPheSerGly	509
DB	2038	TCAGAGAACAGGTCACGGGAAGAGCTGGGGGAAGATGGGCGACTCAGTCTAGCTTTTCGGGC	2097
QY	510	SerMetGluIleIleGluValSer	517

[illegible]



ACCESSION AB052156  
 VERSION AB052156.1 GI:3548676  
 KEYWORDS  
 SOURCE  
 ORGANISM Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 Masuda,K., Shima,H., Watanabe,M. and Kikuchi,K.  
 MKP-7, a novel mitogen-activated protein kinase phosphatase,  
 functions as a shuttle protein  
 J. Biol. Chem. 276 (42), 39002-39011 (2001)  
 JOURNAL MEDLINE 21486429  
 PUBMED 11489891  
 REFERENCE 2 (bases 1 to 3521)  
 AUTHORS Masuda,K., Shima,H. and Kikuchi,K.  
 TITLE Direct Submission  
 JOURNAL Submitted (05-DEC-2000) Kohei Masuda, Institute for Genetic  
 Medicine, Division of Biochemical Oncology and Immunology, Kita-ku-  
 kitsu15-jou nishi7-tyoumei, Sapporo, Hokkaido 060-0815, Japan  
 (E-mail:kou@imm.hokudai.ac.jp, Tel.81-11-706-5536,  
 Fax:81-11-707-6839)  
 FEATURES  
 source  
 location/Qualifiers  
 1..3521  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /chromosome="12"  
 /cell\_line="Jurkat"  
 1..3521  
 /gene="MKP-7"  
 564..2561  
 /gene="MKP-7"  
 /codon\_start=1  
 /product="WAPK phosphatase-7"  
 /protein\_id="BAB40814.1"  
 /db\_xref="GI:13548677"  
 /translation="MAHEMIGTQIVTERLVALLSEGTKEKVLIIIDSRPFVEYNTSHILE  
 AINISCKLMKRLLQDQKVLITELIQSAHKHVDICSKVYVQSSQDVASLSDC  
 FLVLVGLKLFNSVHLLAGGFBSRCFPGLCBKSTLTPCTISQCLPVAINGPT  
 RILPNLVLCQRDVLANKEQNGICVNLASNTCPKDFIPESHFLRVPNDSPCEK  
 KLPDSDSDFPIEKASNGCVLVHCLAGISRSATIAIAYINKRMDSLDEARFVKE  
 KRPTISNDFPGQLDIYEKIKTKGASGPASKLHLHLEKPNPPVAVSEGGQKE  
 TPLSPKCADSTAEAGQRPVHPASVPSPVSPLEDDFLVQALGHLHLSADRLD  
 SNKLKPLFDIISVYSASMAASLHFGSSDALEYKPTTLDGINKLCSFVQGE  
 LSEQTPSPDKBEAIPKKLTQARPSQSKRLSHSVRTSSGTAQSLSLPLHRSQS  
 VEDNHTSFLFLGSLTQOHLTKSAGLGLKWHSDILAQTPTSLTSSWTFATESHFH  
 YSASAIYGGASASVNSQSLPTCGDQVYVVRCKPQSDRADSRSSWHESSPFPEKQF  
 RRSQMFEGSISMSNRSRRELKGVGSQSFSFGSMEILEVS"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 0 Length: 3521  
 Score: 488.00 Matches: 488  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 104.39% Indels: 0  
 DB: 9 Gaps: 0  
 US-09-964-277-21 (1-517) x AB052156 (1-3521)  
 QY 30 GluLeuMetGlnGlnAenGlyIleGlyTyrValLeuAsnAlaSerAsnThrCysProIys 49  
 Db 1095 GAGCTGATGACGACGAATGGATGGTTGGTTATGTTAAATGCCAGCAATACCTGTGCCAAG 1154  
 QY 50 ProAspPheIleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGlu 69  
 Db 1155 CQTGACTTTATCCCGAGTCTCATTTCTCGGTGTCCTGTGAATGACAGCTTTTGTGAG 1214  
 QY 70 LysIleLeuProTyrLeuAspLysSerValAspPheIleGluLysAlaLysAlaSerAsn 89  
 Db 1215 AAAATTTTGGCGTGGTTGGCAAACTAGTAGTATTCATTGAGAAAGCAAAAGCCTCCAAT 1274



```
Db 2355 CCAGCGCGCAGACCCAGGTGACAGAGCTGCTCGCGCGGAGCTGGCATGAAGAGAGC 2414
Qy 470 ProPheGluLysGlnPheLysArgSerCysGlnMetGluPheGlyGluSerIleMet 489
Db 2415 CCCTTGAAGAGCAGTTAAACGCGAGAGCTGCCAAATGGAATTTGGAGAGAGCATCATG 2474
Qy 490 SerGluAsnArgSerArgGluGluGluGlyLysValGlySerGlnSerSerPheSerGly 509
Db 2475 TCAGAGACAGTCCACGGGAGAGCTGGGGAAGTGGGACAGTCAGTCTAGCTCTTTTCGGGC 2534
Qy 510 SerMetGluIleGluValSer 517
Db 2535 AGCATGGAATCATTTAGAGTCTCC 2558

RESULT 11
AX260340 3544 bp DNA linear PAT 26-OCT-2001
LOCUS AX260340
DEFINITION Sequence 1 from Patent WO0173059.
ACCESSION AX260340
VERSION AX260340.1 GI:16509303
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Meyers, R.A.
TITLE 38692 and 21117: dual specificity phosphatase molecules and uses
therefor
JOURNAL Patent: WO 0173059-A 1 04-OCT-2001;
Millennium Pharmaceuticals, Inc. (US)
FEATURES
source Location/Qualifiers
1..3544
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
589..2586
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD10218.1"
/db_xref="GI:16509304"
/db_xref="REMBEL:CAD10218"
/translation="MAHEMIGTQIVTERLVALLSGETKVLIDSRPFVEYNTSHILE
AININCSKMRRLQDDQKVLITELIQHSAKHQDIDCSQKVYVDQSSQVSLSDSC
FLVLGLKESFNSVHLGAGFAFCFGLCEGKSTLVPICISQCLPVAINGPT
RILPNLYGCRDVLNKLQMGQVGLVNASNTCPKPDFIPESHFLRVVNDSCFK
ILPWLKSDVDFIEKAKSGCVLVHCLAGISRSATIAIAYIMKMDMSLDEAYRFVKE
KRPTISPNFNLGQLDYKKIKNOTGASGPKSKILLHLEKNEVPVAVSEGQKSE
TELSPPCADSATSEACORPVHPASVPSPVQPSLLEDSPVQALSGHLISADLED
SNLKRFSLDIKSVSYSAWASLHGFSSSEDALEYKSTTIDGINKLCQFSPVQE
LSEQTPETSFDKEASIPKLOTARPDSQSKRLHVSRTSSGTQAQRSLSPHRSQS
VEDNYHTSFLGLSTSQOHLTKSAGLKGWHSIDLAPQTSPLSTSSWYFATESHSF
YSASAIYGGASAYSAYSQSLPTCGDQVYSVRRRQKPSDRADSRSSWHBESPFQKQF
RRSCQMEFGESIMSENRRBELKVGQSQSFSGSMELIEVS"

ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 3544
Score: 488.00 Matches: 488
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 94.39% Indels: 0
DB: 6 Gaps: 0

US-09-964-277-21 (1-517) x AX260340 (1-3544)

Qy 30 GluLeuMetGlnGlnAnsglyIleGlyTyrValLeuAsnAlaSerAsnThrCysProLys 49
Db 1120 GAGCTGATGCACAGATGGGATGGTTATGTTGTTAAATGCCAGCAATACCTGTCCAAAG 1179
Qy 50 ProAppPheIleProGluSerHisPheLeuArgValProValAsnAppSerPheCysGlu 69
```

```
Db 1180 CCTGACTTTATCCCGAGTCTCATTTCTCTCGTGTGCTGTGAATGACAGCTTTTCTGAG 1239
Qy 70 LysIleLeuProTrpLeuAspLysSerValAspPheIleGluLysAlaLysAlaSerAsn 89
Db 1240 AAAATTTTGGCGTGGTGGCAAAATCAGTAGATTCTATTGAGAAAGCAAAAGCCTCAAT 1299
Qy 90 GlyCysValLeuValHisCysLysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAla 109
Db 1300 GGATGTGTTCTAGTGTCACTGTTTAGCTGGGATCTCCGCTCCGCCACCATCGTATCGCC 1359
Qy 110 TyrIleMetLysArgMetAspMetSerLeuAspGluAlaTyrArgPheValLysGluLys 129
Db 1360 TACATCATGAAGAGATGGACATGCTCTTTAGATGAAGCTTACAGATTCTGTGAAGAAAAA 1419
Qy 130 ArgProThrIleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGluLysLys 149
Db 1420 AGACCTACTATATCTCAAACTTCAATTTCTGGGCAACTCTCTGGACATATGAGAAGAAG 1479
Qy 150 IleLysAsnGlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuHisLeuGlu 169
Db 1480 ATTAAGAACACGACCTGGAGCATCAGGGCCAAAGAGCAAACTCAAGCTGTGCACCTGGAG 1539
Qy 170 LysProAsnGluProValProAlaValSerGluGlyGlyGlnLysSerGluThrProLeu 189
Db 1540 AAGCCAAATGAACCTGCTCCTCTGTCTCAGAGGGTGGACAGAAAAAGCAGAGCCCTTC 1599
Qy 190 SerProProCysAlaAspSerAlaThrSerGluAlaGlyGlnArgProValHisPro 209
Db 1600 AGTCCACCTGTGCCGACTCTGCTACTCAGAGAGCAGGAGCAAAAGGCCCTGCATCCC 1659
Qy 210 AlaSerValProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuVal 229
Db 1660 GCCAGCGTGGCCACGCTGCCAGCGTGCAGCGCTGCTGTAGAGACAGCCGCTGGTA 1719
Qy 230 GlnAlaLeuSerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnLysLys 249
Db 1720 CAGCGCTCAGTGGCTGCACCTGTCCGACAGAGGCTGGAGACAGCAATAAGCTCAAG 1779
Qy 250 ArgSerPheSerLeuAspIleLysSerValSerTyrSerAlaSerMetAlaLysLeu 269
Db 1780 CGTTCTCTCTCTGGATATCAATCAGTTTCTATTTACGCCAGCATGGCAGCATCTTTA 1839
Qy 270 HisGlyPheSerSerSerGluAspAlaLeuGluTyrTyrLysProSerThrThrLeuAsp 289
Db 1840 CATGCTCTCTCTCATCAGAAATGCTTTGGAATATCAAACTTCCACTACTCTGGAT 1899
Qy 290 GlyThrAsnLysLeuCysGlnPheSerProValGlnLeuSerGluGlnThrProGlu 309
Db 1900 GGGACCAACAGCTATGCCAGTTCTCCCTGTTTCAGAACTATCGGAGCAGACTCCCGAA 1959
Qy 310 ThrSerProAspLysGluGluAlaSerIleProLysLysLeuGlnThrAlaArgProSer 329
Db 1960 ACCAGTCTGTATAAGGAGGAGGCCAGCATCCCCAAGAGCTGCAGACCCCGAGCCCTTCA 2019
Qy 330 AspSerGlnSerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArg 349
Db 2020 GACAGCCAGACGACGATGTCATTGCTCGTCCAGAACCCAGCAGCATGGCACCCCGAGAG 2079
Qy 350 SerLeuLeuSerProLeuHisArgSerGlySerValGluAspAsnTyrHisThrSerPhe 369
Db 2080 TCCCTTTTATCTCCACTGTCATCGAAGTGGGAGCGTGGAGACAATTTACCACACCACTTC 2139
Qy 370 LeuPheGlyLeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLeuLys 389
Db 2140 CTTTTCGGCTTTCCACACCCAGCAGCAGCTCACGAGTCTGCTGGCTGGGCTTAAG 2199
Qy 390 GlyTrpHisSerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTrp 409
Db 2200 GGCTGGCACTCGGATATCTTGGCCCCCAGACCTCTACCCCTTCCCTGACCCAGCAGCTGG 2259
Qy 410 TyrPheAlaThrGluSerSerHisPheTyrSerAlaSerAlaIleTyrGlyGlySerAla 429
Db 2260 TATTTGCCACAGAGTCTCTACACTTCTACTCTGCTCTACCCCATCTACGAGGAGCAGTGGC 2319
```





Db 2004 ACCAGTCCTGATAGGAGGAGCCAGCATCCCAAGAGCTGCAGACGCCAGCCCTTCA 2063  
Qy 330 AspSerGlnSerIysArgLeuHisSerValArgThrSerSerGlyThrAlaGlnArg 349  
Db 2064 GACAGCCAGAGCAAGCGATTGCGTTCAGATCGGTTCAGAACAGCAGCAGTGGCAGCCAGAGG 2123  
Qy 350 SerLeuLeuSerProLeuHisArgSerGlySerValGluAspAsnTyrHisThrSerPhe 369  
Db 2124 TCCCTTTTATCTCCACTGATCGAATGGAGTGGAGCGTGGAGGACATACACACAGCTTC 2183  
Qy 370 LeuPheGlyLeuSerThrSerGlnGlnHisLeuThrIysSerAlaGlyLeuGlyLeuLys 389  
Db 2184 CTTTTCGGGCTTTCCACACAGCAGCAGCAGCTTCACGAAGTCTGCTGGGCTTTAAG 2243  
Qy 390 GlyTyrHisSerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTyr 409  
Db 2244 GGTGTCATCTCGATATCTTGGCCCCCAGACCTCTACCCCTTCCCTGACAGCAGCTGG 2303  
Qy 410 TyrPheAlaThrGluSerSerHisPheTyrSerAlaSerAlaIleTyrGlySerAla 429  
Db 2304 TATTTTGGCACAGAGTCTCACACTTCTACTCTGCTCAGCCATCTACGGAGGAGTGCC 2363  
Qy 430 SerTyrSerAlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerVal 449  
Db 2364 AGTACTCTGCTTACAGTCSAGCAGCTGCCACTTCCGGAGACCAAGTCTATCTGTG 2423  
Qy 450 ArgArgGlnGlyProSerAspArgAlaAspSerArgAsgSerTyrHisGluLeuSer 469  
Db 2424 CGCAGCGCGCAGAGCAAGTGCAGAGCTGACTCGCGCGGAGCTGGCATGAGAGAGC 2483  
Qy 470 ProPheGlyGlnGlnPheIysArgSerCysGlnMetGluPheGlyGluSerIleMet 489  
Db 2484 CCTTTTGAAGAAGCAGTTTAAACGCGAAGAGTGCCTCAAAATGGAATTTGGAGAGAGCATG 2543  
Qy 490 SerGluAsnArgSerArgGluGluLeuGlyLysValGlySerGlnSerSerPheSerGly 509  
Db 2544 TCAGAGAAACAGTCCAGGAGAGGTGGGAAAGTGGCAAGTGGCAAGTCTAGCTTTTCGGC 2603  
Qy 510 SerMetGluIleGluValSer 517  
Db 2604 AGCATGGAATCATTTGAGGTCTCC 2627

RESULT 13  
AX374994  
LOCUS AX374994 3766 bp DNA linear PAT 01-MAR-2002  
DEFINITION Sequence 17 from Patent WO210363.  
ACCESSION AX374994  
VERSION AX374994.1 GI:19169826  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Tang Y.T., Elliott, V.S., Ramkumar, J., Yao, M.G., Burford, N., Wang Y.E., Stewart, E.A., Gandhi, A.R., Patterson, C., Lee, E.A., Hafalia, A.J., Lu, D.A., Tribouley, C.M., Griffin, J.A., Baughn, M.R., Yue, H., Warren, B.A., Nguyen, D.B. and Wallia, N.K.  
TITLE Protein phosphatases  
JOURNAL Patent: WO 0210363-A 17 07-FEB-2002;  
Incyte Genomics, Inc. (US)  
FEATURES  
source  
1..3766  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
/note="Incyte ID No: 7480570CB1"

ORIGIN  
Alignment Scores:  
Pred. No.: 0 Length: 3766  
Score: 488.00 Matches: 488  
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 94.39% Indels: 0  
DB: 6 Gaps: 0  
US-09-964-277-21 (1-517) x AX374994 (1-3766)  
Qy 30 GluLeuMetGlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerAsnThrCysProLys 49  
Db 1069 GAGCTGATGCAGCAGATGGGATTTGGTTATGTGTTAAATGCCAGCAATACCTGTCTCCAAAG 1128  
Qy 50 ProAspPheIleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGlu 69  
Db 1129 CTTGACTTTATCCCGAGTCTCATTTCTGCTGCTGCTGTGAATGACAGCTTTGTGAG 1188  
Qy 70 LysIleLeuProTyrLeuAspLysSerValAspPheIleGlyLysAlaLysAsn 89  
Db 1189 AAAATTTTGGCGTGGTGGCAAAATCAGTAGATTTCATTGAGAAAGCAAAAGCTCCCAAT 1248  
Qy 90 GlyCysValLeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAla 109  
Db 1249 GGATGTGTTCTAGTGCACTGTTTAGCTGGGATCTCCCGCTCCGCCACCATCGCTATCGCC 1308  
Qy 110 TyrIleMetLysArgMetAspMetSerLeuAspGluAlaTyrArgPheValLysGluLys 129  
Db 1309 TACATCATGAGAGATGGACATGCTTTTAGTAGAAGCTTACAGATTGTGAAAGAAAAA 1368  
Qy 130 ArgProThrIleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGluLysLys 149  
Db 1369 AGACCTACTATATCTCCAAACTTCAATTTCTGGCCAACTCTCTGGACTATGAGAGAG 1428  
Qy 150 IleLysAsnGlnThrGlyAlaSerGlyProLysSerLysLysLysLeuLysLeuGlu 169  
Db 1429 ATTAAAGAACAGACTGGAGCATCAGGGCCAAAGAGCAAACTCAAGCTGTGCACCTGGAG 1488  
Qy 170 LysProAsnGluProValProAlaValSerGluGlyGlnLysSerGluThrProLeu 189  
Db 1489 AAGCCAAATGAACCTGTCTCTGCTGTCTCAGAGGTGGACAGAAAGCAGAGAGCCCTC 1548  
Qy 190 SerProCysAlaAspSerAlaThrSerGluAlaAlaGlyGlnArgProValHisPro 209  
Db 1549 AGTCACCTCTGTGCCGACTCTGCTACCTCAGAGCAGCAGCAAAAGGCCGCTGCATCCC 1608  
Qy 210 AlaSerValProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuVal 229  
Db 1609 GCCAGCGTCCCGAGCGTCCCGAGCGTGCCTGTTAGAGGACAGCCCGCTGGTA 1668  
Qy 230 GlnAlaLeuSerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnLysLeuLys 249  
Db 1669 CAGGCGCTCAGTGGCTGCACTCTGCCAGACAGGCTGGAGAGACAGCAATAAGCTCAAG 1728  
Qy 250 ArgSerPheSerLeuAspIleLysSerValSerTyrSerAlaSerMetAlaAlaSerLeu 269  
Db 1729 CGTTCCTTCTCTGGATATCAAAATCAGTTTATTCAGCCAGCATGGCAGCATCTTTA 1788  
Qy 270 HisGlyPheSerSerSerGluAspAlaLeuGluTyrTyrLysProSerThrThrLeuAsp 289  
Db 1789 CATGCTTCTCTCATCAGAGATGCTTTGGATACTACAACTTCCACTCTCTGGAT 1848  
Qy 290 GlyThrAsnLysLeuCysGlnPheSerProValGlnGluLeuSerGluGlnThrProGlu 309  
Db 1849 GGGACCAACAGCTATGCCAGTTCTCCCTGTTCCAGAACTATCGGAGCAGATCCCGAA 1908  
Qy 310 ThrSerProAspLysGluAlaSerIleProLysLysLeuGlnThrAlaArgProSer 329  
Db 1909 ACCAGTCTCTGATAGGAGGAGCCAGCATCCCNAGAGGCTGCAGACCCGCGCCTTCA 1968  
Qy 330 AspSerGlnSerLysArgLeuHisSerValArgThrSerSerGlyThrAlaGlnArg 349  
Db 1969 GACAGCCAGAGCAAGCGATTGCTTCGTCAGAACCCAGCAGCAGTGGCACCGCCAGAGG 2028  
Qy 350 SerLeuLeuSerProLeuHisArgSerGlySerValGluAspAsnTyrHisThrSerPhe 369  
Db 2029 TCCCTTTTATCTTCCATGCTCATGGAAGTGGGAGCGTGGAGCAATATTACACACAGCTTC 2088

```
QY 370 LeuPheGlyLeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLeuLys 389
DB 2089 CTTTTGGCCCTTTCACAGCAGCAGCAGCCTCACGAAGTCTGCTGGCCCTTAAG 2148
QY 390 GlyTrpHisSerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTrp 409
DB 2149 GGCTGGCACTCGGATATCTTGGCCCCCAGACCTTACCCCTCCCTGACCAAGCAGCTGG 2208
QY 410 TyrPheAlaThrGluSerSerHisPheTyrSerAlaSerAlaIleTyrGlyGlySerAla 429
DB 2209 TATTTTGGCACAGAGTCTCTACACTTCTACTCTGCTCAGCCATCTACGAGGAGCAGTGGC 2268
QY 430 SerTyrSerAlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerVal 449
DB 2269 AGTTACTCTGCTACAGCTGGAGCCAGCTGCCACCTTGGGAGACCAAGTCTATTCTGTG 2328
QY 450 ArgArgGlnGlnLysProSerAspAlaAspSerArgArgSerTrpHisGluGluSer 469
DB 2329 CGCAGCGGCAGAACCAAGTGCAGAGCTGACTCGCGCGGAGCTGGCATGAAGAGAGC 2388
QY 470 ProPheGluLysGlnPheLysArgArgSerCysGlnMetGluPheGlyGluSerIleMet 489
DB 2389 CCCTTTGAAGACGAGTTAAACGCAGAGAGCTGCCAAATGGAATTTGGAGAGACATCATG 2448
QY 490 SerGluAsnArgSerArgGluGluLeuGlyLysValGlySerGlnSerSerPheSerGly 509
DB 2449 TCAGAGAACAGGTACGGGAGAGCTGGGNAAGTGGGCACTCAGTCTAGCTTTTCGGGC 2508
QY 510 SerMetGluIleIleGluValSer 517
DB 2509 AGCATGGAATCATTTAGGTCTCC 2532
```

```
RESULT 14
BD171157
LOCUS BD171157 4790 bp DNA linear PAT 17-JAN-2003
DEFINITION Novel gene and protein encoded thereby.
ACCESSION BD171157
VERSION BD171157.1 GI:27876969
KEYWORDS WO 02052005-A/13.
```

```
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 4790)
Chara,O., Nagase,T. and Nakajima,D.
Novel gene and protein encoded thereby
Patent: WO 02052005-A 13 04-JUL-2002;
KAZUSA DNA RESEARCH INSTITUTE,OSAMU OHARA,TAKAHIRO NAGASE, DAISUKE
NAKAJIMA
```

```
COMMENT
OS Homo sapiens (human)
FN WO 02052005-A/13
PD 04-JUL-2002
PF 20-DEC-2001 WO 2001JP011217
PR 22-DEC-2000 JP 00P 389742
PI OSAMU OHARA,TAKAHIRO NAGASE,DAISUKE NAKAJIMA
PC C12N15/12,C07K14/47
CC Novel gene and protein encoded thereby
FH Key Location/Qualifiers
FT CDS (184)..(2178).
```

```
FEATURES
source
1..4790
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
```

```
ORIGIN
Alignment Scores:
Pred No.: 0 Length: 4790
Score: 488.00 Matches: 488
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 94.39% Indels: 0
```

```
DB: 6 Gaps: 0
US-09-964-277-21 (1-517) x BD171157 (1-4790)
QY 30 GluLeuMetGlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerAsnThrCysProLys 49
DB 715 GAGCTGATGACAGAGAATGGGATTGGTTATGTTAAATGCCAGCAATACCTGTCCAAAG 774
QY 50 ProAspPheIleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGlu 69
DB 775 CCTGACTTTATCCCGAGTCTCATTTCTCGCGTGCCTGTGAATGACAGCTTTTGTGAG 834
QY 70 LysIleLeuProTrpLeuAspLysSerValAspPheIleGluLysAlaLysAlaSerAsn 89
DB 835 AAATTTTGGCGTGGTGGACAAATCAGTAGATTTCATTGAGNAAGCAAAAGCCTCAAT 894
QY 90 GlyCysValLeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAla 109
DB 895 GGATGTCTTCTAGTGCACATGTTTAGCTGGGATCTCCCGCTCCGCCACCATCGTATCGCC 954
QY 110 TyrIleMetLysArgMetAspMetSerLeuAspGluAlaTyrArgPheValLysGluLys 129
DB 955 TACATCATGAAGAGGATGGACATGTTCTTAGATGAAGCTTACAGATTGTGAAGAAAAA 1014
QY 130 ArgProThrIleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGluLysLys 149
DB 1015 AGACCTACTATATCTCAAACTTCAATTTCTGGGCCAACTCTCGACTACTGAGAAGAG 1074
QY 150 IleLysAsnGlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuHisLeuGlu 169
DB 1075 ATTAAGAACAGAGCTGGAGCATCAGGGCCAAAGAGCAAACTCAAGCTGTGCACCTGGAG 1134
QY 170 LysProAsnGluProValProAlaValSerGluGlyGlnLysSerGluThrProLeu 189
DB 1135 AGCCAAATGAACCTGCTCCTCTGCTCTCAGAGGTGGACAGAAAGCCAGACGCCCTTC 1194
QY 190 SerProCysAlaAspSerAlaThrSerGluAlaAlaGlyGlnArgProValHisPro 209
DB 1195 AGTCCACCTGTGCGCACTCTGCTAGCTCAGAGGCAGCAGGCAAAAGGCCGTGATCC 1254
QY 210 AlaSerValProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuVal 229
DB 1255 GCCAGCTGCCAGCGTGGCCAGCTGCGAGCTGCTGTTAGAGCAGCCGCTGGTA 1314
QY 230 GlnAlaLeuSerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnLysLeuLys 249
DB 1315 CAGCGCTCAGTGGGTGCACTGTCCGACAGACAGGCTGGAAGACAGCAATAAGCTCAAG 1374
QY 250 ArgSerPheSerLeuAspIleLysSerValSerTyrSerAlaSerMetAlaIleSerLeu 269
DB 1375 CGTTCTCTCTCTGGATATCAAAATCAGTTTCATATTCAGCCAGCATGCGACATCCCTTA 1434
QY 270 HisGlyPheSerSerGluAspAlaLeuGluTyrTyrLysProSerThrThrLeuAsp 289
DB 1435 CATGGCTTCTCTCATCAGAAGATGCTTTGGAATACTACAAACCTTCCACTACTCTGGAT 1494
QY 290 GlyThrAsnLysLeuCysGlnPheSerProValGlnGluLeuSerGluGlnThrProGlu 309
DB 1495 GGGACCAACAGAGTATGCCAGTTCTCCCTGTTCAGGAACTATCGAGCAGACTCCCCGAA 1554
QY 310 ThrSerProAspLysGluGluAlaSerIleProLysLysLeuGlnThrAlaArgProSer 329
DB 1555 ACCAGTCTCTATAGGAGGAGGAGCAGCATCCCCAGAGAGCTGCAGCCGAGGCTTCA 1614
QY 330 AspSerGlnSerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArg 349
DB 1615 GACAGCCAGAGCAAGCGATTGCAATTCGGTCAGAACCCAGCAGCAGTGGCACCCGCCAGAGG 1674
QY 350 SerLeuLeuSerProLeuHisArgSerGlySerValGluAspAsnTyrHisThrSerPhe 369
DB 1675 TCCCTTTTATCTCACTGATCATCGAAGTGGAGCGGTGGAGACAATTTACCACACAGCTTC 1734
QY 370 LeuPheGlyLeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLeuLys 389
```

```

Db      1735  CTTTTCGGCTTCCACGACGACGACGACCTCAGAACTGCTGGCTGGCTTAAAG 1794
Qy      390  GlyTyrPheSerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTyr 409
Db      1795  GGCTGGCACTCGGATATCTTGGCCCCCAGAGACTCTACCCCTTCCCTGACGAGAGCTGG 1854
Qy      410  TyrPheAlaThrGluSerSerHisPheTyrSerAlaSerAlaIleTyrGlySerAla 429
Db      1855  TATTTTGCACAGAGTCTTCACTTCTACTCTGCTCAGCCATCTACGAGGAGTGCC 1914
Qy      430  SerTyrSerAlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerVal 449
Db      1915  AGTTACTTGCCTTACAGCTGCAGCCAGCTGCCACTTGGCGAGACCAAGTCTATTCTGTG 1974
Qy      450  ArgArgGlnLysProSerAspArgAlaAspSerArgArgSerTyrPheGluGluSer 469
Db      1975  CGAGGCGGAGAGGCAAGTACAGAGCTGACTCGCGCGAGCTGGCATGAAGAGAGC 2034
Qy      470  ProPheGluLysGlnPheLysArgArgSerCysGlnMetGluPheGlyGluSerIleMet 489
Db      2035  CCCTTTGAAAGCAGTTTAAACGACGAGCTGCCAAATGGAATTTGGAGAGAGCATCATG 2094
Qy      490  SerGluAsnArgSerArgGluGluLeuGlyLysValGlySerGlnSerSerPheSerGly 509
Db      2095  TCAGAGAACAGAGTTCACGGGAAGAGCTGGGGAAAGTGGCAGTCACTAGCTTTTCGGGC 2154
Qy      510  SerMetGluIleIleGluValSer 517
Db      2155  AGCATGGAATCATTTAGAGTCTCC 2178

RESULT 15
BD183422 4790 bp DNA linear PAT 17-JUN-2003
LOCUS Novel genes and proteins encoded by the genes.
DEFINITION BD183422
ACCESSION BD183422
VERSION BD183422.1 GI:31875622
KEYWORDS JP 2002345492-A/135.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 4750)
AUTHORS Chara,O., Nagase,T. and Nakajima,D.
TITLE Novel genes and proteins encoded by the genes
JOURNAL Patent: JP 2002345492-A 135 03-DEC-2002;
KAZUSA DNA RESEARCH INSTITUTE
COMMENT OS Homo sapiens (human)
PN JP 2002345492-A/135
PD 03-DEC-2002
PF 26-FEB-2002 JP 2002049009
PI OSAMU OHARA, TAKAHIRO NAGASE, DAISUKE NAKAJIMA
PC C12N15/09,C07K14/47//A61K31/711,A61K38/00,A61K48/00,A61P25/00,
PC A61P25/14,
PC A61P25/18,A61P35/00,C12N15/00,A61K37/02
CC Novel genes and proteins encoded by the genes FH Key
CDS Location/Qualifiers
FT CDS Location/Qualifiers
source 1..4790
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 0 Length: 4790
Score: 488.00 Matches: 488
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 94.39% Indels: 0
DB: 6 Gaps: 0

```

```

US-09-964-277-21 (1-517) x BD183422 (1-4790)
Qy      30  GluLeuMetGlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerAsnThrCysProLys 49
Db      715  GAGCTGATGTCAGCAGATGGGATGGTTATGTGTAAATGCCAGCAATACCTGTCCAAAG 774
Qy      50  ProAspPheIleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGlu 69
Db      775  CTTGACTTTATCCCGAGTCTCATTTCTGCTGCTGCTGTGAATGACAGCTTTGTGAG 834
Qy      70  LysIleLeuProTyrPheLeuAspLysSerValAspPheIleGluLysAlaLysAlaSerAsn 89
Db      835  AAAATTTTGGCGTGGTTGGCAAAATCAGTAGATTTCATTGAGAAAGCAAAAGCCTCCAAT 894
Qy      90  GlyCysValLeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAla 109
Db      895  GGATGTGTTCTAGTCACTGTTTAGCTGGGATCTCCCGCTCCGCCACATCGCTATCGCC 954
Qy      110  TyrIleMetLysArgMetAspMetSerLeuAspGluAlaTyrArgPheValLysGluLys 129
Db      955  TACATCATGAGAGGATGGACATGTCTTTAGATGAAGCTTACAGATTGTGAAAGAAAA 1014
Qy      130  ArgProThrIleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGluLysLys 149
Db      1015  AGACCTACTATATCTCCAAACTTCAATTTCTGGCCAACTCTCTGGACTATGAGAGAGAG 1074
Qy      150  IleLysAsnGlnThrGlyValAserGlyProLysSerLysLysLeuLysLeuHisLeuGlu 169
Db      1075  ATTAAGAACCCAGACTGGAGCATCAGGCCCAAGAGCAAACTCAAGCTGCTGCACCTCGAG 1134
Qy      170  LysProAsnGluProValProAlaValSerGluGlyGlyGlnLysSerGluThrProLeu 189
Db      1135  AAGCCAAATGAACCTGTCTCTCTCAGAGGGTGGACAGAAAAAGCAGAGCCCTC 1194
Qy      190  SerProProCysAlaAspSerAlaThrSerGluAlaAlaGlyGlnArgProValHisPro 209
Db      1195  AGTCCACCTGTGCGGCTGCTGCTGCTCAGAGCAGCAGGACAAAGCCCGTGCATCCC 1254
Qy      210  AlaSerValProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuVal 229
Db      1255  GCCAGCTGCCAGCGTCCCAGCGGTGCAGCGCTGCTGTAGAGGACAGCCCGCTGGTA 1314
Qy      230  GlnAlaLeuSerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnLysLysLys 249
Db      1315  CAGGCGCTCAGTGGGCTGCACTGTCCGACAGAGGCTGGAGACAGCAATAAGCTCAAG 1374
Qy      250  ArgSerPheSerLeuAspIleLysSerValSerTyrSerAlaSerMetAlaAlaSerLeu 269
Db      1375  CGTTCTCTCTCTGATATCAATCAGTTTCATATTACCCAGCATGGCAGCATCCTTA 1434
Qy      270  HisGlyPheSerSerSerGluAspAlaLeuGluTyrTyrLysProSerThrThrLeuAsp 289
Db      1435  CATGCGCTTCTCTCATCAGAAGATGCTTTGGAATPACTACAAACCTTCCACTACTCTGGAT 1494
Qy      290  GlyThrAsnLysLeuCysGlnPheSerProValGlnGlnLeuSerGluGlnThrProGlu 309
Db      1495  GGGACCAACAGCATATGCCACTTCTCCCTGTTCCAGGAATACTATCGGAGCAGACTCCCGAA 1554
Qy      310  ThrSerProAspLysGluGluAlaSerIleProLysLysLeuGlnThrAlaArgProSer 329
Db      1555  ACCAGTCTCTGATAAGAGAGAAAGCCAGCATTCCTCCCAAGAGCTGCAGACCCGCGCTTCA 1614
Qy      330  AspSerGlnSerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArg 349
Db      1615  GACGCCAGACCAAGCGATTGCAATTCGTCAGAACCCAGCAGCAGTGGCACCGCCGAGGG 1674
Qy      350  SerLeuLeuSerProLeuHisArgSerGlySerValGluAspAsnTyrHisThrSerPhe 369
Db      1675  TCCCTTTTATCTCCATGCTATCGAAGTGGAGCGTGGAGGACAATTACCACACAGCTTC 1734
Qy      370  LeuPheGlyLeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuGlyLys 389
Db      1735  CTTTTCGGCCCTTTCCACCAGCAGCAGCACTCTACGAAAGTCTGCTGGCCTTGGCCCTTAAAG 1794

```

```
QY 390 GlyTrpHisSerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTrp 409
Db 1795 GGCTGGCACTCGGATATCTTTGGCCCCCAGACCTCTACCCCTTCCTGACCAGCAGCTGG 1854
QY 410 TyrPheAlaThrGluSerSerHisPheTyrSerAlaSerAlaIleTyrGlyGlySerAla 429
Db 1855 TATTTTCCACAGAGTCTCTACACTTCTACTCTGCTCAGCCATCTACGGAGGCGAGTGCC 1914
QY 430 SerTyrSerAlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerVal 449
Db 1915 AGTTACTCTGCTACAGCTGCAGCCAGCTGCCCACTTGGCGAGACCAAGTCTATTCTGTG 1974
QY 450 ArgArgArgGlnLysProSerAspArgAlaAspSerArgArgSerTyrHisGluGluSer 469
Db 1975 CGCAGCGCGCAGAGCCAGTGCAGAGCTGACTCGCGCGAGCTGGCATGGAGAGAGC 2034
QY 470 ProPheGluLysGlnPheLysArgArgSerCysGlnMetGluPheGlyGluSerIleMet 489
Db 2035 CCCTTTGAAAAGCAGTTTAAACGCAGAGCTGCCAAATGGAATTTGGAGAGAGCATCATG 2094
QY 490 SerGluAsnArgSerArgGluGluLeuGlyLysValGlySerGlnSerSerPheSerGly 509
Db 2095 TCAGAGAACAGGTCAACGGGAAGAGCTGGGAAAGTGGGCAGTCAGTCTAGCTTTTCGGGC 2154
QY 510 SerMetGluIleIleGluValSer 517
Db 2155 AGCATGGAAATCATTTGAGGTCTCC 2178
```

Search completed: February 28, 2004, 03:08:22  
Job time : 6077 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model  
Run on: February 28, 2004, 00:57:54 ; Search time 110 seconds  
(without alignments)  
2608.269 Million cell updates/sec

Title: US-09-964-277-21  
Perfect score: 517  
Sequence: 1 MLPLSLQTVFLYFWNRR.....LGKVGQSFGSGMEIIIVS 517

Scoring table:  
Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Zgapop 6.0 , Zgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues  
Word size: 1

Total number of hits satisfying chosen parameters: 1360453

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlh  
-O=/cgn2\_1/USPTO.spool/US09964277/runat 25022004 085402 2228/app query.fasta 1.711  
-DB=Issued Patents NA -QWTF=fastcap -SUFFIX=olip2n.rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZES=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09964277 @CGN 1 1 56 @runat 25022004 085402 2228 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Issued Patents NA.\*  
1: /cgn2\_6/ptodata/2/ina/5A COMB.seq.\*  
2: /cgn2\_6/ptodata/2/ina/5B COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/6A COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/6B COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/PTUS COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	488	94.4	1998	US-09-816-494-3	Sequence 3, Appli
2	480	94.4	2544	US-09-816-494-1	Sequence 1, Appli
3	20	3.9	2377	US-09-920-668-3	Sequence 3, Appli
4	12	2.3	2109	US-09-016-434-1135	Sequence 1135, App
5	12	2.3	2109	US-09-023-655-946	Sequence 946, App
6	9	1.7	320	US-09-389-681-310	Sequence 310, App
7	9	1.7	320	US-09-620-405B-310	Sequence 310, App
8	9	1.7	320	US-09-339-338-310	Sequence 310, App
9	9	1.7	320	US-09-433-826B-310	Sequence 310, App
10	9	1.7	320	US-09-604-287A-310	Sequence 310, App
11	9	1.7	320	US-09-834-759-310	Sequence 310, App
12	9	1.7	539	US-09-389-681-311	Sequence 311, App

c 13	9	1.7	539	4	US-09-620-405B-311	Sequence 311, App
c 14	9	1.7	539	4	US-09-339-338-311	Sequence 311, App
c 15	9	1.7	539	4	US-09-433-826B-311	Sequence 311, App
c 16	9	1.7	539	4	US-09-604-287A-311	Sequence 311, App
c 17	9	1.7	539	4	US-09-834-759-311	Sequence 311, App
c 18	9	1.7	944	4	US-09-371-671B-10	Sequence 10, Appl
19	9	1.7	1238	2	US-08-530-290-11	Sequence 11, Appl
20	9	1.7	1238	4	US-09-702-705-803	Sequence 803, App
21	9	1.7	1238	4	US-09-736-457-803	Sequence 803, App
22	9	1.7	1238	4	US-09-614-124B-803	Sequence 803, App
23	9	1.7	1238	4	US-09-671-325-803	Sequence 803, App
24	9	1.7	1238	4	US-09-589-184-803	Sequence 803, App
25	9	1.7	1619	4	US-09-702-705-801	Sequence 801, App
26	9	1.7	1619	4	US-09-736-457-801	Sequence 801, App
27	9	1.7	1619	4	US-09-614-124B-801	Sequence 801, App
28	9	1.7	1619	4	US-09-671-325-801	Sequence 801, App
29	9	1.7	1619	4	US-09-589-184-801	Sequence 801, App
30	9	1.7	1987	2	US-08-990-379-1	Sequence 1, Appli
31	9	1.7	1993	2	US-08-990-379-2	Sequence 2, Appli
32	9	1.7	2000	4	US-09-016-434-1291	Sequence 1291, App
33	9	1.7	2064	4	US-09-702-705-825	Sequence 825, App
34	9	1.7	2064	4	US-09-736-457-825	Sequence 825, App
35	9	1.7	2064	4	US-09-614-124B-825	Sequence 825, App
36	9	1.7	2064	4	US-09-671-325-825	Sequence 825, App
37	9	1.7	2064	4	US-09-589-184-825	Sequence 825, App
38	9	1.7	2109	4	US-09-702-705-826	Sequence 826, App
39	9	1.7	2109	4	US-09-736-457-826	Sequence 826, App
40	9	1.7	2109	4	US-09-614-124B-826	Sequence 826, App
41	9	1.7	2109	4	US-09-671-325-826	Sequence 826, App
42	9	1.7	2109	4	US-09-589-184-826	Sequence 826, App
43	9	1.7	2240	4	US-09-016-434-1100	Sequence 1100, App
44	9	1.7	4637	4	US-09-702-705-804	Sequence 804, App
45	9	1.7	4637	4	US-09-736-457-804	Sequence 804, App

ALIGNMENTS

RESULT 1  
US-09-816-494-3  
; Sequence 3, Application US/09816494  
; Patent No. 6664089  
; GENERAL INFORMATION:  
; APPLICANT: Meyers, Rachel A.  
; TITLE OF INVENTION: 38692 AND 21117, NOVEL DUAL SPECIFICITY  
; FILE REFERENCE: 10448-030002  
; CURRENT APPLICATION NUMBER: US/09/816,494  
; CURRENT FILING DATE: 2001-03-23  
; PRIOR APPLICATION NUMBER: US 60/191,858  
; PRIOR FILING DATE: 2000-03-24  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 1998  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-816-494-3

Alignment Scores:  
Pred. No.: 0 Length: 1998  
Score: 488.00 Matches: 488  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 94.39% Indels: 0  
DB: 4 Gaps: 0

US-09-964-277-21 (1-517) x US-09-816-494-3 (1-1998)

Qy	30	GlutMetGlnGlnAsnGlyVileGlyVileValLeuAsnAlaSerAsnThrCysProLys 49
Db	532	GAGCTGATGACGACGATGGATGGTATGCTTAATGCCAGCAATACCTGCTCCAAAG 591
Qy	50	ProAspPheIleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGlu 69

592	DB	CCTGACTTTATCCCGAGTCTCAATTTCCCTGCGTGTGCTGTGTAATGACAGCTTTTGTGAG	651
70	QY	LysIleLeuProTTPLeuAspLysSerValAspPheIleGluLysAlaLysAlaSerAsn	89
652	DB	AAATTTTGGCGTGTTGGACAAATCAGTAGATTTCATTGAGAAACAAAGCCCTCCAAAT	711
90	QY	GlyCysValLeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAla	109
712	DB	GGATGTGTTCCTAGTGCACCTGTTTAGCTGGGATCTCCCGCTCCGCCACCATCGCTATCGCC	771
110	QY	TyrIleMetLysArgMetAspMetSerLeuAspGluAlaTyrArgPheValLysGluLys	129
772	DB	TACATCATGAAGAGATGGACATGTCTTTAGATGAAGCTTACAGATTGTGAAGAAAAA	831
130	QY	ArgProThrIleSerProAsnPheAsnPheLeuGlyGlnLeuLeuAspTyrGluLys	149
832	DB	AGACCTACTATATCTCAAACTTCAATTTCTGGGCCAACTCCTCGCATATGAGAAGAAG	891
150	QY	IleLysAsnGlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuLeuHisLeuGlu	169
892	DB	ATTAGAACCCAGACTGGAGCAATCAGGGCCAGAGGCAAACTCAGCTCTCGACTCGAG	951
170	QY	LysProAsnGluProValProAlaValSerGluGlyGlnLysSerGluThrProLeu	189
952	DB	AAGCCAAATGAACCTGTCTCTGTCTCAGAGGGTGGACAGAAAGAGAGACGCCCTC	1011
190	QY	SerProProCysAlaAspSerAlaThrSerGluAlaAlaGlyGlnArgProValHisPro	209
1012	DB	AGTCCACCGTGTGGCGACTCTGTACTCTCAGAGCAGCAGGCAAAAGCCCGTGCATCCC	1071
210	QY	AlaSerValProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuVal	229
1072	DB	GCCAGCGTGCCCGGTGCCCGGTGACCGTGCCTGTGTAGAGGACAGCCCGCTGGTA	1131
230	QY	GlnAlaLeuSerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnLysLeuLys	249
1132	DB	CAGCGCTCAGTGGCGTGCACCTGTCCGCAGACAGGCTGGAAAGACAGCAATAAAGCTCAAG	1191
250	QY	ArgSerPheSerLeuAspIleLysSerValSerTyrSerAlaSerMetAlaAlaSerLeu	269
1192	DB	CGTTCCTTCTCTCTGGATATCAATCAGTTTCATATTAGCCAGCATGGCAGCATCTTTA	1251
270	QY	HisGlyPheSerSerSerGluAspAlaLeuGluTyrTyrLysProSerThrThrLeuAsp	289
1252	DB	CATGCTTCTCCTCATCAGAAGATGCTTTGGAATACTACAAACCTTCCACTACTCTGAT	1311
290	QY	GlyThrAsnLysLeuCysGlnPheSerProValGlnGluLeuSerGluGlnThrProGlu	309
1312	DB	GGGACCAACAGCTATGCCAGTTCTCCCTGTTCAGAACTATCGGACGACACTCCCGAA	1371
310	QY	ThrSerProAspLysGluGluAlaSerIleProLysLysLeuGlnThrAlaArgProSer	329
1372	DB	ACCAGTCTCTGATAGGAGGAGCCAGCATCCCAAGAAGCTCAGACCGCCAGGCCCTTCA	1431
330	QY	AspSerGlnSerLysArgLeuHisSerValArgThrSerSerSerGlyThrAlaGlnArg	349
1432	DB	CACAGCCAGACCAAGCATTTGCATTTCCGTTCAGAACCCAGCAGCAGTGGCCCGCCAGAGG	1491
350	QY	SerLeuLeuSerProLeuHisArgSerGlySerValGluAspAsnTyrHisThrSerPhe	369
1492	DB	TCCCTTTTATCTCCACTGTCAATGAGTGGAGCGTGGAGACAATTACCACACAGCTTC	1551
370	QY	LeuPheGlyLeuSerThrSergInGlnHisLeuThrLysSerAlaGlyLeuGlyLys	389
1552	DB	CTTTTCGGGCTTTCCACAGCAGCAGCACCTCAGCAAGTCTGTCTGGCTGGGCTTAAG	1611
390	QY	GlyTTPHisSerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerSerTrp	409
1612	DB	GGCTGGCACTCGGATATCTTGGCCCCCAGACCTCTTACCCCTTCCCTGACCAGCAGCTGG	1671
410	QY	TyrPheAlaThrGluSerSerHisPheTyrSerAlaSerAlaIleTyrGlyLysAla	429

```

Db      1672  TATTTCGCCAGAGTCTCTACACACTTCTACTCTGCTCTAGCCATCTACGGAGGCGAGTGCC 1733
QY      430  SerTyrSerAlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerVal 449
Db      1732  AGTTACTCTGCTACAGCTGCACCGCAGCTGCCACCTTGCAGAGACCAAGCTATTATCTGTG 1791
QY      450  ArgArgArgGlnTyrProSerAspArgAlaAspSerAspGlySerTrpHisGluGluSer 469
Db      1792  CGCAGGCGGAGAGGCAAGCAAGTACAGAGCTGACTCGCGCGAGCTGGCATGAAGAGAGC 1851
QY      470  ProPheGluLysGlnPheLysArgTyrSerCysGlnMetGluPheGlyGluSerIleMet 489
Db      1852  CCCTTTGAAGAAGCAGTTTAAACGCAGAACCTGCCAAATGGAAATTTGGAGAGAGCATCATG 1911
QY      490  SerGluAsnArgSerArgGluLeuGlyLysValGlySerGlnSerSerPheSerGly 509
Db      1912  TCAGAGAACACAGGTACCGGAGAGAGCTGGGGAAAGTGGGCGAGTCAGCTTAGCTTTTCGGGC 1971
QY      510  SerMetGluIleLeuGluValSer 517
Db      1972  AGCATGGAAATCATTTGAGGTCTCC 1995

RESULT 2
US-09-816-494-1
; Sequence 1, Application US/09816494
; Patent No. 6664089
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; TITLE OF INVENTION: 38632 AND 21117, NOVEL DUAL SPECIFICITY
; TITLE OF INVENTION: PHOSPHATASE MOLECULES AND USES THEREFOR
; FILE REFERENCE: 10448-030002
; CURRENT APPLICATION NUMBER: US/09/816,494
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 60/191,858
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3544
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (589) ... (2583)
US-09-816-494-1

Alignment Scores:
Pred. No.: 0 Length: 3544
Score: 488.00 Matches: 488
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 94.39% Indels: 0
DB: 4 Gaps: 0

US-09-964-277-21 (1-517) x US-09-816-494-1 (1-3544)
QY      30  GluLeuMetGlnGlnAsnGlyIleGlyTyrValLeuAsnAlaSerAsnThrCysProLys 49
Db      1120  GAGCTGTATGCAGCAGAAATGGCAATGGTTATGTGTATAATGCCAGCAATACCTGTCCAAAG 1179
QY      50  ProAspPheIleProGluSerHisPheLeuArgValProValAsnAspSerPheCysGlu 69
Db      1180  CTTGACTTATATCCCGAGTCTCATTTCTTGCCTGTGCTCTGATGATGACAGCTTTTGTGAG 1239
QY      70  LysIleLeuProTrpLeuAspLysSerValAspPheIleGluLysAlaLysAlaSerAsn 89
Db      1240  AAAATTTTGGCGTGTGGCAAAATCAGTAGATTTTCATTGAGAAAGCAAAAGCGCTCCAAT 1299
QY      90  GlyCysValLeuValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAla 109
Db      1300  GGATGTGTCTTAGTGCACTGTTTAGTCTGGGATCTCCCGCTCGCCACCATCGCTATCGCC 1359
QY      110  TyrIleMetLysArgMetAspMetSerLeuAspGluAlaTyrArgPheValLysGluLys 129

```

Db 1360 TACATCATGAGGATGGACATGCTTTAGATGAGCTTACAGTTTGTGAAGAAAA 1419  
QY 130 ArgProThrIleSerProAsnPhenPheLeuGlyGlnLeuLeuAspTyrGluLeuLys 149  
Db 1420 AGACCTACTATATCTCCAACTTCAATTTCTGGGCCAACTCTCTGGACTATGAGAAG 1479  
QY 150 IleLysAsnGlnThrGlyAlaSerGlyProLysSerLysLeuLysLeuHisLeuGlu 169  
Db 1480 ATTAGAACACGACTGGAGCATCAGGGCCAAAGACCAACTCAAGCTGCTGACCTGGAG 1539  
QY 170 LysProAsnGluProValProAlaValSerGluGlyGlnLysSerGluThrProLeu 189  
Db 1540 AAGCCAAATGAACCTGTCTCTCTCTCAGAGGTGGACAGAAAGCAGAGCCCTC 1599  
QY 190 SerProProCysAlaAspSerAlaThrSerGluAlaGlyGlnArgProValHisPro 209  
Db 1600 AGTCCACCTGTGGCGACTCTGTCTACCTCAGAGGAGCAGACAAAGCCGCTGATCC 1659  
QY 210 AlaSerValProSerValProSerValGlnProSerLeuLeuGluAspSerProLeuVal 229  
Db 1660 GCCAGCGTGCCAGCGTGCCAGCGTGCGCGTCTGTAGAGGACAGCCCGCTGTA 1719  
QY 230 GlnAlaLeuSerGlyLeuHisLeuSerAlaAspArgLeuGluAspSerAsnLysLeuLys 249  
Db 1720 CAGGCGCTCAGTGGCTGCACCTGTCCGACAGACGCTGGAGACAGCAATAAGCTCAAG 1779  
QY 250 ArgSerPheSerLeuAspIleLysSerValSerTyrSerAlaSerMetAlaAlaSerLeu 269  
Db 1780 CGTTCCTCTCTCTCGATATCAATCAGTTTCATATTACGCCAGCATGGCGATCTCTTA 1839  
QY 270 HisGlyPheSerSerSerGluAlaLeuGluTyrTyrLysProSerThrThrLeuAsp 289  
Db 1840 CATGCTCTCTCTCATCAGAGATGCTTTGGATACTACAAACCTTCCACTCTGGAT 1899  
QY 290 GlyThrAsnLysLeuCysGlnPheSerProValGlnGluLeuSerGluGlnThrProGlu 309  
Db 1900 GGGACCAACAGCTATGCGAGTCTCCCTCTTCCAGAACTATCGGAGCAGATCCCGAA 1959  
QY 310 ThrSerProAspLysGluGluAlaSerIleProLysLysLeuGlnThrAlaArgProSer 329  
Db 1960 ACCAGTCTGTATAGAGGAAGCCAGCATCCCAAGAGCTGCAGACCGCCAGCCCTCA 2019  
QY 330 AspSerGlnSerLysArgLeuHisSerValArgThrSerSerGlyThrAlaGlnArg 349  
Db 2020 GACAGCCAGAGCAAGCGATTGCAATCGGTGAGAACCCAGACGATGCGCCAGAGG 2079  
QY 350 SerLeuLeuSerProLeuHisArgSerGlySerValGluAspAsnTyrHisThrSerPhe 369  
Db 2080 TCCCTTTTATCTCCACTGCATCGAGTGGAGCGTGGAGACATACCAACAGCTTC 2139  
QY 370 LeuPheGlyLeuSerThrSerGlnGlnHisLeuThrLysSerAlaGlyLeuLys 389  
Db 2140 CTTTTGGCCCTTCCACGAGCAGCAGCCTCAGAAAGTCTGCTGGCCCTGGCCCTTAAG 2199  
QY 390 GlyTyrPheHisSerAspIleLeuAlaProGlnThrSerThrProSerLeuThrSerTrp 409  
Db 2200 GCTGCGCACTGGATATCTTGGCCCCCAGACCTCTACCCCTTCCCTGACGACGCTGG 2259  
QY 410 TyrPheAlaThrGluSerSerHisPheTyrSerAlaSerAlaIleTyrGlyGlySerAla 429  
Db 2260 TATTTTGGCACAGATCTCTACATCTCTGCTCAGCCATCTAGCGAGGACGCTGCC 2319  
QY 430 SerTyrSerAlaTyrSerCysSerGlnLeuProThrCysGlyAspGlnValTyrSerVal 449  
Db 2320 AGTTACTCTGCTCAGCTCAGCGACGCTGCCACTTGGCGGAGACCAAGTCTATTCTGTG 2379  
QY 450 ArgArgArgGlnLysProSerAspArgAlaAspSerArgArgSerTrpHisGluGluSer 469  
Db 2380 CGCAGCGGAGAGACCCAGTGCAGAGCTGACTCCGGCGGAGCTGGCATGAGAGAGC 2439  
QY 470 ProPheGluLysGlnPheLysArgArgSerCysGlnMetGluPheGlyGluSerIleMet 489

Db 2440 CCCTTTGAAAGCAGTTTAAACGAGAGAGCTGCCAAATGGAATTGGAGAGAGCATCATG 2499  
QY 490 SerGluAsnArgSerArgGluGluLeuGlyLysValGlySerGlnSerSerPheSerGly 509  
Db 2500 TCAGAGAAACAGCTCAGCGGAAGAGCTGGGAAAGTGGGAGTCTAGCTTTTCGGGC 2559  
QY 510 SerMetGluIleIleGluValSer 517  
Db 2560 AGCATGGAATCATTTAGGTCTCC 2583  
RESULT 3  
US-09-920-668-3  
; Sequence 3, Application US/09920668  
; Patent No. 6482644  
; GENERAL INFORMATION:  
; APPLICANT: Lex M. Cowser  
; APPLICANT: Brett P. Monia  
; TITLE OF INVENTION: ANTISENSE MODULATION OF DUAL SPECIFIC PHOSPHATASE 8 EXPRESSION  
; FILE REFERENCE: RTS-0246  
; CURRENT APPLICATION NUMBER: US/09/920,668  
; CURRENT FILING DATE: 2001-08-01  
; NUMBER OF SEQ ID NOS: 49  
; SEQ ID NO 3  
; LENGTH: 2377  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (135)...(2012)  
US-09-920-668-3  
Alignment Scores:  
Pred. No.: 2,78e-10 Length: 2377  
Score: 20.00 Matches: 20  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 3.87% Indels: 0  
Gaps: 0  
US-09-964-277-21 (1-517) x US-09-920-668-3 (1-2377)  
QY 94 ValHisCysLeuAlaGlyIleSerArgSerAlaThrIleAlaIleAlaTyrIleMetLys 113  
Db 864 GTCCACTGTCTGGTGGCATCTCCCGCTCTGCCACCATGCCCTACATCATGAAG 923  
RESULT 4  
US-09-016-434-1135  
; Sequence 1135, Application US/09016434  
; Patent No. 6500938  
; GENERAL INFORMATION:  
; APPLICANT: Janice Au-Young  
; APPLICANT: Jeffrey J. Seilhamer  
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING  
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION  
; NUMBER OF SEQUENCES: 1490  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3174 PORTER DRIVE  
; CITY: PALO ALTO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/016,434  
; FILING DATE: HERewith  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:



SEQUENCE CHARACTERISTICS:  
LENGTH: 2109 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GENBANK  
CLONE: g1418933  
US-09-023-655-946

Alignment Scores:  
Pred. No.: 0.0302 Length: 2109  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 2.32% Indels: 0  
DB: 4 Gaps: 0

US-09-964-277-21 (1-517) x US-09-016-434-1135 (1-2109)

Qy 92 ValLeuValHisCysLeuAlaGlyIleSerArgSer 103  
Db 1216 GTCTTGGTACATTGCTTGGCTGCATTAGCCGCTCA 1251

RESULT 5  
US-09-023-655-946  
; Sequence 946, Application US/09023655  
; Patent No. 6607879  
; GENERAL INFORMATION:  
; APPLICANT: Cocks, Benjamin G.  
; APPLICANT: Susan G. Stuart  
; APPLICANT: Jeffrey J. Seilhamer  
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE  
; TITLE OF INVENTION: EXPRESSION  
; NUMBER OF SEQUENCES: 1508  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3174 PORTER DRIVE  
; CITY: PALO ALTO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/023,655  
; FILING DATE: HEREWITH  
; CLASSIFICATION:  
; PRIOR APPLICATION NUMBER:  
; APPLICATION NUMBER:  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Zeller, Karen J.  
; REGISTRATION NUMBER: 37,071  
; REFERENCE/DOCKET NUMBER: PA-0001 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (650) 845-0555  
; TELEFAX: (650) 845-4166  
; INFORMATION FOR SEQ ID NO: 946:

SEQUENCE CHARACTERISTICS:  
LENGTH: 2109 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GENBANK  
CLONE: g1418933  
US-09-023-655-946

Alignment Scores:  
Pred. No.: 0.0302 Length: 2109  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 2.32% Indels: 0  
DB: 4 Gaps: 0

US-09-964-277-21 (1-517) x US-09-023-655-946 (1-2109)

Qy 92 ValLeuValHisCysLeuAlaGlyIleSerArgSer 103  
Db 1216 GTCTTGGTACATTGCTTGGCTGCATTAGCCGCTCA 1251

RESULT 6  
US-09-389-681-310/c  
; Sequence 310, Application US/09389681A  
; Patent No. 6518237  
; GENERAL INFORMATION:  
; APPLICANT: Yuqiu, Jiang  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Xu, Jiangchun  
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND  
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE  
; FILE REFERENCE: 210121.470C3  
; CURRENT APPLICATION NUMBER: US/09/389,681A  
; CURRENT FILING DATE: 1999-09-02  
; NUMBER OF SEQ ID NOS: 463  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 310  
; LENGTH: 320  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: (250)  
; OTHER INFORMATION: n=A,T,C or G  
US-09-389-681-310

Alignment Scores:  
Pred. No.: 4.74 Length: 320  
Score: 9.00 Matches: 9  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.74% Indels: 0  
DB: 4 Gaps: 0

US-09-964-277-21 (1-517) x US-09-389-681-310 (1-320)

Qy 98 AlaGlyIleSerArgSerAlaThrIle 106  
Db 74 GCAGGCAATTCCTCCGTCAGCCACCATC 48

RESULT 7  
US-09-620-405B-310/c  
; Sequence 310, Application US/09620405B  
; Patent No. 6528054  
; GENERAL INFORMATION:  
; APPLICANT: Jiang, Yuqiu  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Xu, Jiangchun

APPLICANT: Harlocker, Susan L.  
APPLICANT: Hepler, William T.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER

FILE REFERENCE: 210121.470C8

CURRENT APPLICATION NUMBER: US/09/620,405B

CURRENT FILING DATE: 2000-07-20

NUMBER OF SEQ ID NOS: 495

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 310

LENGTH: 320

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc feature

LOCATION: (250)

OTHER INFORMATION: n=A,T,C or G

US-09-620-405B-310

Alignment Scores:  
Pred. No.: 4.74 Length: 320  
Score: 9.00 Matches: 9  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.74% Indels: 0  
DB: 4 Gaps: 0

US-09-964-277-21 (1-517) x US-09-620-405B-310 (1-320)

QY 98 AlaGlyIleSerArgSerAlaThrIle 106

Db 74 GCAGGCATTTCCTCCGTCAGCCACCATC 48

## RESULT 8

US-09-339-338-310/c

Sequence 310, Application US/09339338A

Patent No. 6573368

GENERAL INFORMATION:

APPLICANT: Yuqiu, Jiang

APPLICANT: Dillon, Davin C.

APPLICANT: Mitcham, Jennifer L.

APPLICANT: Xu, Jiangchun

TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND

TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE

FILE REFERENCE: 210121.470C2

CURRENT APPLICATION NUMBER: US/09/339,338A

CURRENT FILING DATE: 1999-06-23

NUMBER OF SEQ ID NOS: 315

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 310

LENGTH: 320

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc feature

LOCATION: (250)

OTHER INFORMATION: n=A,T,C or G

US-09-339-338-310

Alignment Scores:  
Pred. No.: 4.74 Length: 320  
Score: 9.00 Matches: 9  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.74% Indels: 0  
DB: 4 Gaps: 0

US-09-964-277-21 (1-517) x US-09-339-338-310 (1-320)

QY 98 AlaGlyIleSerArgSerAlaThrIle 106

Db 74 GCAGGCATTTCCTCCGTCAGCCACCATC 48

## RESULT 9

US-09-433-826B-310/c

Sequence 310, Application US/09433826B

Patent No. 6579973

GENERAL INFORMATION:

APPLICANT: Jiang, Yuqiu

APPLICANT: Dillon, Davin C.

APPLICANT: Mitcham, Jennifer L.

APPLICANT: Xu, Jiangchun

APPLICANT: Harlocker, Susan L.

TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND

TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE

FILE REFERENCE: 210121.470C4

CURRENT APPLICATION NUMBER: US/09/433,826B

CURRENT FILING DATE: 1999-11-03

NUMBER OF SEQ ID NOS: 474

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 310

LENGTH: 320

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc feature

LOCATION: (250)

OTHER INFORMATION: n=A,T,C or G

US-09-433-826B-310

Alignment Scores:  
Pred. No.: 4.74 Length: 320  
Score: 9.00 Matches: 9  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.74% Indels: 0  
DB: 4 Gaps: 0

US-09-964-277-21 (1-517) x US-09-433-826B-310 (1-320)

QY 98 AlaGlyIleSerArgSerAlaThrIle 106

Db 74 GCAGGCATTTCCTCCGTCAGCCACCATC 48

## RESULT 10

US-09-604-287A-310/c

Sequence 310, Application US/09604287A

Patent No. 6586572

GENERAL INFORMATION:

APPLICANT: Jiang, Yuqiu

APPLICANT: Dillon, Davin C.

APPLICANT: Mitcham, Jennifer L.

APPLICANT: Xu, Jiangchun

APPLICANT: Harlocker, Susan L.

APPLICANT: Hepler, William T.

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER

FILE REFERENCE: 210121.470C7

CURRENT APPLICATION NUMBER: US/09/604,287A

CURRENT FILING DATE: 2000-06-22

NUMBER OF SEQ ID NOS: 489

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 310

LENGTH: 320

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc feature

LOCATION: (250)

OTHER INFORMATION: n=A,T,C or G

US-09-604-287A-310

Alignment Scores:  
Pred. No.: 4.74 Length: 320  
Score: 9.00 Matches: 9  
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.74% Indels: 0  
DB: 4 Gaps: 0

US-09-964-277-21 (1-517) x US-09-604-287A-310 (1-320)

Qy 98 AlaGlyIleSerArgSerAlaThrIle 106

Db 74 GCAGGCATTTCCCGGTCCAGCCACCATC 48

## RESULT 11

US-09-834-759-310/c  
; Sequence 310, Application US/09834759  
; Patent No. 6860197  
; GENERAL INFORMATION:  
; APPLICANT: Jiang, Yugu  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Hepler, William T.  
; APPLICANT: Henderson, Robert A.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER  
; FILE REFERENCE: 210121.470C9  
; CURRENT APPLICATION NUMBER: US/09/834,759  
; CURRENT FILING DATE: 2001-04-13  
; NUMBER OF SEQ ID NOS: 547  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 310  
; LENGTH: 320  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: (250)  
; OTHER INFORMATION: n=A,T,C or G  
US-09-834-759-310

Alignment Scores:  
Pred. No.: 4.74 Length: 320  
Score: 9.00 Matches: 9  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.74% Indels: 0  
DB: 4 Gaps: 0

US-09-964-277-21 (1-517) x US-09-834-759-310 (1-320)

Qy 98 AlaGlyIleSerArgSerAlaThrIle 106

Db 74 GCAGGCATTTCCCGGTCCAGCCACCATC 48

## RESULT 12

US-09-389-681-311/c  
; Sequence 311, Application US/09389681A  
; Patent No. 6518237  
; GENERAL INFORMATION:  
; APPLICANT: Yuqui, Jiang  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Xu, Jiangchun  
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND  
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE  
; FILE REFERENCE: 210121.470C3  
; CURRENT APPLICATION NUMBER: US/09/389,681A  
; CURRENT FILING DATE: 1999-09-02  
; NUMBER OF SEQ ID NOS: 463  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 311  
; LENGTH: 539  
; TYPE: DNA  
; ORGANISM: Homo sapiens

US-09-389-681-311

Alignment Scores:  
Pred. No.: 8.08 Length: 539  
Score: 9.00 Matches: 9  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.74% Indels: 0  
DB: 4 Gaps: 0

US-09-964-277-21 (1-517) x US-09-389-681-311 (1-539)

Qy 98 AlaGlyIleSerArgSerAlaThrIle 106

Db 136 GCAGGCATTTCCCGGTCCAGCCACCATC 110

## RESULT 13

US-09-620-405B-311/c  
; Sequence 311, Application US/09620405B  
; Patent No. 6528054  
; GENERAL INFORMATION:  
; APPLICANT: Jiang, Yugu  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Hepler, William T.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER  
; FILE REFERENCE: 210121.470C8  
; CURRENT APPLICATION NUMBER: US/09/620,405B  
; CURRENT FILING DATE: 2000-07-20  
; NUMBER OF SEQ ID NOS: 495  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 311  
; LENGTH: 539  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-620-405B-311

Alignment Scores:  
Pred. No.: 8.08 Length: 539  
Score: 9.00 Matches: 9  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.74% Indels: 0  
DB: 4 Gaps: 0

US-09-964-277-21 (1-517) x US-09-620-405B-311 (1-539)

Qy 98 AlaGlyIleSerArgSerAlaThrIle 106

Db 136 GCAGGCATTTCCCGGTCCAGCCACCATC 110

## RESULT 14

US-09-339-338-311/c  
; Sequence 311, Application US/09339338A  
; Patent No. 6573368  
; GENERAL INFORMATION:  
; APPLICANT: Yuqui, Jiang  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Xu, Jiangchun  
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND  
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE  
; FILE REFERENCE: 210121.470C2  
; CURRENT APPLICATION NUMBER: US/09/339,338A  
; CURRENT FILING DATE: 1999-06-23  
; NUMBER OF SEQ ID NOS: 315  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 311  
; LENGTH: 539  
; TYPE: DNA

```

; ORGANISM: Homo sapiens
US-09-339-338-311

Alignment Scores:
Pred. No.:      8.08      Length:      539
Score:          9.00      Matches:      9
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      1.74%      Indels: 0
DB:               4      Gaps: 0

US-09-964-277-21 (1-517) x US-09-339-338-311 (1-539)

QY 98 AlaGlylleSerArgSerAlaThrIle 106
DB 136 GCAGGCATTCCCGGTGAGCCACCATC 110

RESULT 15
US-09-433-826B-311/c
; Sequence 311, Application US/09433826B
; Patent No. 6579973
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqui
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; FILE REFERENCE: 210121.470C4
; CURRENT APPLICATION NUMBER: US/09/433.826B
; CURRENT FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 474
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 311
; LENGTH: 539
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-433-826B-311

Alignment Scores:
Pred. No.:      8.08      Length:      539
Score:          9.00      Matches:      9
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      1.74%      Indels: 0
DB:               4      Gaps: 0

US-09-964-277-21 (1-517) x US-09-433-826B-311 (1-539)

QY 98 AlaGlylleSerArgSerAlaThrIle 106
DB 136 GCAGGCATTCCCGGTGAGCCACCATC 110

Search completed: February 28, 2004, 04:14:26
Job time : 117 secs

```